DOCUMENT RESUME

ED 473 169 CG 031 104

AUTHOR	McCutcheon, Lynn E., Ed.
TITLE	North American Journal of Psychology, 2001.
ISSN	ISSN-1527-7143
PUB DATE	2001-00-00
NOTE	549p.; Published semi-annually.
AVAILABLE FROM	NAJP, 240 Harbor Dr., Winter Garden, FL 34787 ($30 per annual subscription). Tel: 407-877-8364; Web site: http://najp.8m.com.
PUB TYPE	Collected Works - Serials (022)
JOURNAL CIT	North American Journal of Psychology; v3 n1-3 2001
EDRS PRICE	EDRS Price MF02/PC22 Plus Postage.
DESCRIPTORS	*Psychology; *Research Tools; *Scholarly Journals; *Social Science Research

ABSTRACT

"North American Journal of Psychology" publishes scientific papers of general interest to psychologists and other social scientists. Articles included in volume 3 issue 1 (March/April 2001) are: "Sense of Humor in Black and White"; "Convergent Validity of the Situational Outlook Questionnaire"; "Alcohol Consumption and Consequences in a Sample of University Undergraduates"; "Psychology and Religion: Indicators of Integration"; "Descriptions of Religious Experience among U.S. Christians and Malaysian Muslims"; "The Best Laid Plans of Mice and Men: The Role of Decision Confidence in Outcome Success"; "Disposition toward Thinking Critically: A Comparison of Preservice Teachers and Other University Students"; "An Interview with Robert Sternberg about Learning Disabilities"; "Interdisciplinarity and the Cross-Training of Clinical Psychologists: Preparing Graduates for Hybrid Careers"; "Validity Comparison of the General Ability Measure for Adults with the Wonderlic Personnel Test." Articles included in volume 3 issue 2 (June/July 2001) are: "Parental Disciplinary History and Current Level of Empathy and Moral Reasoning among Late Adolescents"; "An Interview with Russell Eisenman"; "Differential Processes of Emotion Space over Time"; "Healthy Perfectionism and Positive Expectations about Counseling"; "Attitudes toward Perpetration of Sexual Abuse Against Children"; "Self-Protective Pessimism: Optimistic Bias in Reverse"; "Lost Letter Technique: Returned Responses to Battered and Abused Women, Men, and Lesbians"; "Stability of Self-Esteem as Moderator of the Relationship between Level of Self-Esteem and Depression"; "Life after College: Psychology Students" Perceptions of Salary, Business Hiring Criteria, and Graduate Admission Criteria." Articles included in volume 3 issue 3 (December 2001) are: "Boundaries in the Mind: Past Research and Future Directions"; "Religion"s Role in the Terroristic Attack of September 11, 2001"; "Adolescent Substance Abuse: An Empirical Examination of the Covalent Security Attribution Model of User and Nonuser Groups"; "Conjunctive Forms and Conditional Inference in Questions and Statements"; "Monitoring Children"s Behavior in a Remote Community before and Six Years after the Availability of Broadcast TV"; "Effects of a Health and Relationship Education Program on Drug Behaviors"; "Attribution and Reciprocity on International Relations: The Attribution Reciprocity Model"; "Self-Definition in Psychotherapy: Is It Time to Revisit Self-Perception Theory?" (GCP)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 473 169

North American Journal of Psychology
Volume 3
2001

Lynn McCutcheon, Editor

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

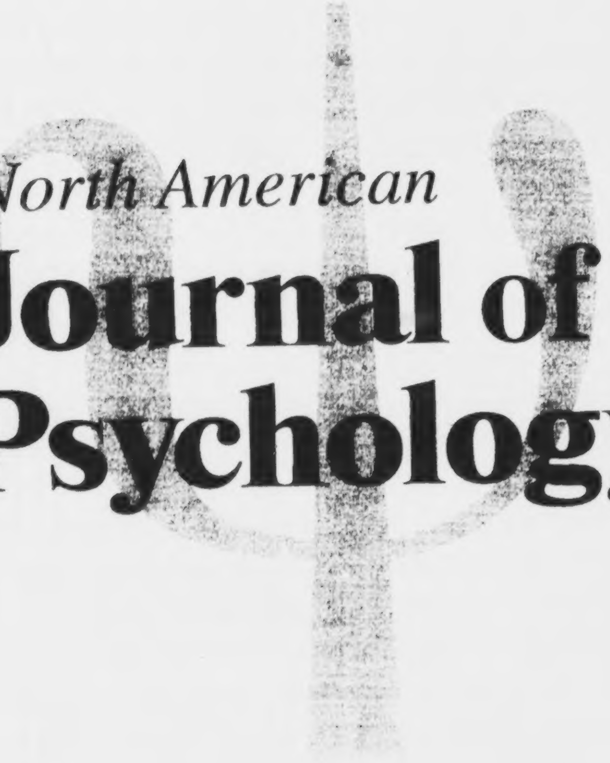
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

McCUTCHEON

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
1

CG031104

BEST COPY AVAILABLE



North American

Journal of Psychology

BEST COPY AVAILABLE

Vol. 3, No. 1 2001

Advisory Editorial Board

ADRIAN FURNHAM
Psychology
University College London

JUTTA HECKHAUSEN
Max Planck Institute for Human
Development and Education

DOUGLAS JACKSON
Psychology
University of Western Ontario

HEIDI KELLER
Psychology
University of Osnabrueck

MANUEL LONDON
Psychology
SUNY, Stony Brook

KAREN LYNESS
Management
Baruch College, CUNY

ROBERT PASNAK
Psychology
George Mason University

RICHARD RYAN
Psychology
University of Rochester

PHILIP ZIMBARDO
Psychology
Stanford University

Editor
LYNN E. McCUTCHEON

How to Subscribe to The *North American* Journal of Psychology

All subscriptions are on a calendar year basis only. If you order late in the year, the earlier issue will be sent to you. Institutional rate: $65. Individual rate: $30. Add $7 for Canadian subscriptions. Add $8 for subscriptions outside North America. *North American* Journal of Psychology is published semi-annually in June and December.

Now visit our website at http://najp.8m.com

North American Journal of Psychology

March/April 2001 **ISSN 1527-7143** Vol. 3, No.1

Sense of Humor in Black and White 1
James A. Thorson, F. C. Powell, & V. T. Samuel

A Comparative Study of Multi-component Weight Loss Treatments: A Replication & Exploratory Look at Female Smokers, Nonsmokers, & Participant BMIs 13
David L. Johnson & Gary D. Brinker

Convergent Validity of the Situational Outlook Questionnaire: Discriminating Levels of Perceived Support for Creativity 31
Scott G. Isaksen & Kenneth J. Lauer

Alcohol Consumption & Consequences in a Sample of University Undergraduates: Patterns & Relations to Internalized Distress & Religious Affiliation 41
J. M. Oliver, Cynthia K. S. Reed, & Bruce W. Smith

Psychology & Religion: Indicators of Integration 61
Amber Haque

Descriptions of Religious Experience Among U.S. Christians & Malaysian Muslims 77
Michael E. Nielsen & Rick Stevens

The Best Laid Plans of Mice & Men: The Role of Decision Confidence in Outcome Success 91
R. Prasad Bingi, David L. Turnipseed, & George M. Kasper

Disposition Toward Thinking Critically: A Comparison of Pre-service Teachers & Other University Students 109
M.E.Eigenberger, K.A Sealander, J.A. Jacobs, & S.M. Shellady

Age-discrepant Relationships: Do These Romances Fare Well? 119
E. Zak, E. Armer, K. Edmunds, T. Fleury, M. Sarris, & B. Shatynski

The Cognitive Style Index and the Legal Profession in Nova Scotia: A Factor Analytic Study 123
H.J. Murphy, P.A. Doucette, W.E. Kelleher, J.G. Reid, & J.D. Young

An Interview With Robert Sternberg About Learning Disabilities 131
Robert Sternberg & Michael F. Shaughnessy

Interdisciplinarity and the Cross-training of Clinical Psychologists: Preparing Graduates for Hybrid Careers 139
Daniel Holland

Book Review: *Guiding Your Child Through Grief* 171
Ross Robak

Validity Comparison of the General Ability Measure for Adults with the Wonderlic Personnel Test 173
J.P. Leverett, T.D. Matthews, K.S. Lassiter, & N.L. Bell

Editor's Comments

This is the first issue for 2001, our third volume. NAJP was issued twice a year for the first two years, but this year we are doing three - this one, which you are probably getting in late March or early April, a summer issue, followed by one in the late Fall.

Thanks again to all the reviewers, all of whom did a marvelous job of critiquing papers promptly and with great insight. Without our reviewers, this journal would simply not have the quality it has today.

As some of you already know, we are now abstracted in Psychological Abstracts. You probably do not need to be reminded how much that means in terms of widespread circulation of the articles we publish. I received a call recently from a person representing a university library. He wanted an article that had appeared in our last issue - found out about it through the online version of Psych. Abstracts. This means that all of our previous issues are or will be abstracted soon.

With this issue, James Jones, one of the most respected individuals in the area of social psychology, joins our advisory editorial board. We are privileged to have him with us. Congratulations to Phil Zimbardo, another board member, for the national honor which he won.

Special thanks to Mike Shaughnessy, who has a knack for conducting great interviews on behalf of NAJP. This issue features an interview with Robert Sternberg, in which he discusses his research and his views on learning disabilities.

Lynn E. McCutcheon, editor

Sense of Humor in Black and White

James A. Thorson
F. C. Powell
University of Nebraska at Omaha

V. T. Samuel
Grambling State University

There are many articles in the literature comparing examples of black and white humor, but there is surprisingly little information providing data giving comparisons of black and white samples in terms of *sense* of humor. This article compares scale and item scores of 116 Black and 357 White university students on the Multidimensional Sense of Humor Scale. Overall scores were not significantly different, and there were only a few significant differences in individual item scores.

"Humor is a way to keep from killing yourself...it removes anxiety. Other kids threw rocks. I make jokes." -- Abe Burrows, as cited in Dorinson & Boskin, 1988.

Ways to describe cultures are found in ethnic humor. Especially with regard to coping humor, some cultures exhibit characteristically self-deprecating humor. This may in fact be a display of strength, to paradoxically poke fun at one's own weakness. Freud (1916) would have argued that this is the highest form of coping. Valliant (1977) calls humor the most elegant of the defenses. Lefcourt and Martin (1986) have demonstrated that humor is used as a potent coping mechanism.

Sense of humor itself has been shown in a number of studies to be related to several beneficial constructs, including creativity (Humke & Schaefer, 1996); coping with stress (Kuiper, McKenzie, & Belanger, 1995); a higher discomfort threshold (Hudak, Dale, & Hudak, 1991); level of intimacy (Hampes, 1994); lower levels of depression (Thorson & Powell, 1994); cheerfulness (Kohler & Ruch, 1996); and relieving tension in difficult circumstances (Rust & Goldstein, 1989; Squire, 1995; vanWormer & Boes, 1997). It may contribute to longevity (Yoder & Haude, 1995) and it is related to both age and gender (Thorson, Powell, Sarmany- Schuller, & Hampes, 1997). It also correlates significantly with the personality traits of exhibition and dominance (Thorson & Powell. 1993b).

Author info: Correspondence should be sent to: James A. Thorson, Dept. of Gerontology, U. of Nebraska at Omaha, Omaha, NE 68182.
Jthorson@mail.unomaha.edu

North American Journal of Psychology, Vol. 3, No. 1, 1-12.

© NAJP

Hardly any sense of humor research has tested for racial differences; indeed, hardly any has had enough African American research subjects to make any racial comparisons. An exception is a study by Smeltzer and Leap (1988), who asked 118 White and 47 Black employees to rate the appropriateness of neutral, sexist, and racist jokes in the workplace. Surprisingly, Whites rated racist jokes as more inappropriate than did Blacks, although differences were slight. This research, however, did not test for overall sense of humor, just for opinions on the appropriateness of particular types of jokes. More usual is the type of article written by Aznam (1989), which points out that humor is a social lubricant that allows people of different races to get along better by laughing at themselves, but without any data presentation or analysis.

One would be hard-pressed to say that there is a characteristically "black" coping humor; one might argue that the African American population is so culturally diverse as to be difficult to compare it to smaller, more socially cohesive groups in this regard. That is, while many Blacks no doubt use coping humor on a daily basis, it would be as improper to draw this as a racial or cultural stereotype for African Americans as it would be to do so for Americans of European heritage. Rather, one might argue that "black jokes" are currently more likely to be told by racist Whites than they are by Blacks. On the other hand, "Jewish jokes," by comparison, are as likely to be told by Jews as by anyone else, and it is a mark of cultural acceptance and assimilation that they often are seen in a less sensitive light than are jokes that are at the expense of African Americans.

An illustration might be the story of the priest and the rabbi who are neighbors. When he gets a new car, the priest blesses it with holy water. Seeing this, the rabbi is at a loss as to what to do when he gets his new car. Thinking that he must do something, he goes out and cuts an inch off the tail pipe. This is the kind of joke that can be (and is) told by Jew and non-Jew alike, with no hard feelings.

Ziv (1984) gives the example of three Jews who have been placed in front of the firing squad. The officer asks the first if he wants a blindfold. "Yes," says the first one; the officer asks the second one if he wants a blindfold, and he says yes. Asked if he wants a blindfold, the third one turns it down, at which point the second Jew leans over and whispers: "Moishe, stop making trouble" (p. 60). This is a classic example of humor in which the participants are coping.

More sensitive would be the joke about the African American who moves into a white neighborhood. Seen cutting the grass, a white neighbor asks him what he charges to mow a lawn. The reply: "I get to sleep with the lady of the house." While it is healthy to laugh at our prejudices, some

Blacks and Whites alike might take exception to such a jest; racial sensitivity is such in both cultures as to make joking less acceptable.

At any rate, we would argue that social norms have changed in recent decades to such a degree that racial humor has lost the easy acceptance it might have found in, say, the 1970s, just as in contemporary society one seldom hears the kind of racist humor that might have been acceptable in the 1950s. So, we might advance the idea that while some cultures can be described by their humor, one would be hard-pressed to say that this is true among African Americans, and that cultural norms about humor vary along with other social changes.

The beauty of humor, however, is in its complexity. Jokes have different meanings to different ears. Take, for example, the story told in the 1960s of the white boater who was seen in Louisiana pulling a black water skier. An observer allowed that it was wonderful to see such unexpected harmony between the races. A good old boy sitting on the dock corrected him: "Naw, Jake's just trolling for gators." Who is the butt of the joke here? It is joking at the expense of the racist White, but it might equally be interpreted as racist humor at the expense of the hapless Black on the water skis. Again, we would argue that this is the kind of anecdote that has lost its social acceptability. It may have been a refreshing way to laugh at racism when the joke was current circa 1968; few - White or Black - would now be caught telling such a story in polite society. Cultural norms change like the seasons, and propriety changes as well. Still, we cannot conclude that there is a typically self-deprecating vein that is characteristic in African American humor similar to, say, Jewish humor.

Apte (1987) observed that ethnic humor has become controversial, particularly when groups are made the butt of humor initiated by anyone but themselves. As a result there is a growing ambivalence toward its use, particularly in the public domain. He went on to observe that a *sense* of humor is a core cultural value in American society. He argued that societal consensus in America is that a person with a sense of humor is more sociable, easier to get along with and to work with, innovative, and capable of overcoming adversities. Thus sense of humor is usually regarded as a virtuous personality trait (1987). Apte then noted, however, the increasing emphasis on cultural and ethnic pluralism, and wondered if assimilation into one dominant culture now has the importance that it once did. "During the melting pot era of emphasis on assimilation, the numbers of ethnic minority groups had to grin and bear the humor of which they were the butt. In a sense they conformed, at least outwardly, to the dominant value of having a sense of humor since they learned to tolerate and, on occasion, to force laughter at the disparaging humor initiated by the outsiders at their expense" (1987, p. 33). Racial jokes have subsequently lost their social sanction.

Early in the 1970s it was observed that no one laughed at racist humor any more. Thus, according to Apte, cultural norms have evolved, and the question has arisen: do groups differ in basic sense of humor, which seems unlikely, or only in what is culturally acceptable as topical humor? "It is interesting that the practice of telling self-disparaging jokes has been common among some ethnic groups such as Jews and Blacks. It appears that such jokes, when told by members of the ethnic groups themselves, are acceptable. Only when outsiders tell them does the joke become public and not tolerated" (1987, p. 38).

Apte concludes that uses of ethnic humor in contemporary society have come under severe constraints, and that members of various ethnic groups no longer are forced to display their sense of humor by laughing at jokes of which they are the target. This says nothing, however, about which groups are higher or lower in basic *sense* of humor, or in what ways various groups might construe humor differently. Apte presents no data or experimental investigation to bolster his arguments.

Similarly, Jaret (1999) has written about attitudes toward ethnic humor. He did a data-based study, surveying 727 Blacks and Whites on their feelings toward the uses of ethnic humor. There were a few inter-racial differences, but the overall conclusion was that Black and White attitudes are not polarized. Jaret, however, did not probe for any racial differences in sense of humor or constructions of humor.

Brabant and Mooney (1999) took the innovative approach of comparing the comic strip *Curtis*, the protagonist and author of which are African American, with similar comics about little boys who happen to be white, *Dennis the Menace* and *Calvin and Hobbes*. While maintaining that Sunday comics reflect cultural values, Brabant and Mooney in fact demonstrated few differences in cultural values. Neither did they attempt any analysis of humor or sense of humor reflected in the respective comic strips.

Davies (1990) argues that ethnic put-down humor is universal throughout the world, and that it would be impossible to find a type of humor that is totally unique to any nationality. Bowles (1994), on the other hand, maintains that there is a distinctly "black" humor that is used as a defense mechanism in response to perceptions of a racist majority culture. This coping humor releases repressed feelings and promotes group cohesiveness. That is, we make fun of that which threatens us.

Brake pointed out in 1975 that what was then referred to as "Negro humor" had started to change from self-effacing, survival humor to a more aggressive anti-racism style of humor. Perhaps, like Jewish humor, there has been less of a need with black humor to "shuffle" in recent years; as social norms between the races have changed in the past several decades, we might hypothesize that types of humor used by both Black and White persons have become more homogeneous. By extension, we might advance for the hypothesis of the present study that there are few actual *sense* of humor differences between whites and blacks.

METHOD

Participants

Samples of 357 White and 116 Black university students completed the Multidimensional Sense of Humor Scale (MSHS). Respondents in the Black sample were students at Grambling State University; they ranged in age from 18 to 51 years (*M* = 21.9 years, *SD* = 4.7). The Whites were students at the University of Nebraska at Omaha. They ranged in age from 17 to 62 (*M* = 24.3 years, *SD* = 8.0). In both instances the university students were volunteers. Asking them to complete the MSHS had previously been approved by the respective Institutional Review Boards.

Instrument and Procedure

The MSHS (Thorson & Powell, 1993a) contains 24 self-descriptive Likert items testing for humor creativity ("Sometimes I think up jokes and funny stories"), coping ("Uses of humor help to put me at ease"), attitudes toward humorous people ("I appreciate those who generate humor"), attitudes toward humor itself ("I like a good joke"), and facility at social uses of humor ("I can ease a tense situation by saying something funny"). Respondents are asked to indicate choices on each item on a five-point scale ranging from "strongly agree" to "strongly disagree." Scoring is from 0 to 4 on each item, with blanks scored as 2. There are 18 positively-phrased items; the scale's six negatively-phrased items are reversed in scoring. In every instance a higher numerical score indicates higher sense of humor. Scores on the MSHS have been shown (Thorson & Powell, 1993a) to vary by gender, males generally scoring higher in humor creativity and females having higher coping humor scores, so comparisons are reported separately for females and males. All participants were given the MSHS at their respective universities, for which they received extra credit in an undergraduate course.

RESULTS

Cronbach alphas for the two groups were .89 for the Black respondents and .91 for White respondents. The .01 level of significance was used because of the large number of comparisons. The total MSHS score for Black women was slightly higher than that for White women, but the difference was not statistically significant. Black women, however, scored significantly higher on two of the 24 scale items:

Sometimes I think up jokes and funny stories.
I'm confident that I can make other people laugh.

The White female respondents scored significantly higher on just two items:

I appreciate those who generate humor.
Uses of humor help to put me at ease.

The Black and White male respondents also did not differ significantly in total MSHS scores; there were significant differences between the two groups on only one of the items. Black respondents were higher on:

Uses of humor help to put me at ease.

There were no statistically significant differences in total MSHS score between Black women ($M = 68.2$, $SD = 12.5$) and Black men ($M = 65.3$, $SD = 10.6$) $t = 1.02$. Similarly, the total scale scores for White males ($M = 67.7$, $SD = 14.3$) and White females ($M = 67.2$, $SD = 12.5$), did not differ significantly ($t = 0.35$).

DISCUSSION

The sense of humor scores for both groups were surprisingly similar, particularly in view of the fact that the Black sample's data was gathered in Louisiana and the White sample's in Nebraska. One might have anticipated at least a few cultural differences based on geography, if not upon race. The male groups in particular were so similar as to merit little further analysis. The total MSHS scores for the White and the Black males were not significantly different, and there was only one of the 24 item scores that displayed statistical difference. Perhaps greater differences might have emerged with larger samples, particularly of African American males. Nevertheless, the item scores indicate no clear pattern of difference. Sometimes total scale scores that are similar emerge that way because of one group being higher in a particular element of humor and the other group being higher in another, the differences balancing each other out. This is not the case with the male respondents in the present study. African Americans were higher on one coping humor item. There is no clear pattern of difference, and we can conclude that, at least within this sample, there is no pronounced sense of humor difference between White and Black males. Geography made no difference, perhaps because all of these respondents have been exposed to national electronic entertainment media virtually all of their lives. There may be subtle *uses* of humor that are race or place specific, but in terms of the elements of *sense* of humor as measured by the Multidimensional Sense of Humor Scale, there are no real differences for all practical purposes. Conclusions more pronounced than that should, of course, be based on multiple studies with large, diverse samples.

A more interesting pattern of difference emerges among the women. The White female respondents indicated higher scores on just one of the coping humor items; the Black women had higher scores on none of the coping humor items. They were not significantly different on the remaining items that tested for coping humor, so perhaps we could call it almost even between the two groups in terms of uses of humor for adaptation or coping.

The White women were higher on one other scale item, the appreciation of humorous people. This is not a clear pattern, but there is some slight evidence here to suggest that white women in this sample place a somewhat higher value on humor appreciation.

The interesting pattern that emerges, however, is with the African American women's higher scores on humor generativity. These respondents say that they think up jokes and funny stories, and are confident that they can make others laugh. Both of these differences were statistically significant.

One of the things that we can say with some certainty about people who generate humor is that they *try* harder. That is, it takes some effort and motivation to create humor out of everyday situations, and it takes a certain uninhibited boldness to venture to share it with others. Other research has found a relationship between humor creativity and the personality trait of exhibition (Thorson & Powell, 1993b). The same study found a relationship between humor creativity and dominance, thus the group control item. The humor creator takes the initiative in the group interaction; she may entertain, certainly, but she also seeks to be in charge and set the agenda.

It might be said that all human interaction is cultural to a certain degree, and uses of humor certainly could be said to be culturally influenced. In this study we have presented some data to indicate that African American women describe themselves as higher in the creativity element of sense of humor, at least in comparison to the sample of white women tested.

This may be something of a hypothesis-building finding. It should be replicated, certainly, but additional research might be in order to determine if younger Black women are less inhibited in uses of humor, if there are cultural patterns that demand greater humor creation efforts on their part, or if there is a kind of humor of oppression in play here that necessitates humor generation as a group survival technique.

REFERENCES

Apte, M. L. (1987). Ethnic humor versus "sense of humor." *American Behavioral Scientist, 30*, 27-41.

Aznam, S. (1989). Quipping away at racism. *Far Eastern Economic Review, 146* (50), 93.

Bowles, D. D. (1994). Black humor as self-affirmation. *Journal of Multicultural Social Work, 3* (2), 1-9.

Brabant, S., & Mooney, L. A. (1999). The social construction of family life in the Sunday comics: Race as a consideration. *Journal of Comparative Family Studies, 30*(1), 113-133.

Brake, R. (1975). The lion act is over: Passive/aggressive patterns of communication in American Negro humor. *Journal of Popular Culture, 9*, 549-559.

Davies, C. (1990). *Ethnic humor around the world*. Bloomington, IN: Indiana University Press.

Dorinson, J., & Boskin, J. (1988). Racial and ethnic humor. In Lawrence E. Mintz (Ed.), *Humor in America: A research guide to genres and topics*. New York: Greenwood press, p. 168.

Freud, S. (1916). *Wit and its relation to the unconscious*. New York: Moffat & Yard.

Hampes, W. P. (1994). Relation between intimacy and the Multidimensional Sense of Humor Scale. *Psychological Reports, 74*, 1360-1362.

Hudak, D. A., Dale, A. J., Hudak, M. A., & DeGood, D. E. (1991). Effects of humorous stimuli and sense of humor on discomfort. *Psychological Reports, 69*, 779-786.

Humke, C., & Schaefer, C. E. (1996). Sense of humor and creativity. *Psychological Reports, 82*, 544-546.

Jaret, C. (1999). Attitudes of whites and blacks toward ethnic humor: A comparison. *Humor: International Journal of Humor Research, 12*(4), 385-409.

Kohler, G., & Ruch, W. (1996). Sources of variance in current sense of humor inventories: How much substance, how much method variance? *Humor: International Journal of Humor Research, 9*, 363-397.

Kuiper, N. A., McKenzie, S. D., & Belanger, K. A. (1995). Cognitive appraisals and individual differences in sense of humor: Motivational and affective implications. *Personality and Individual Differences, 19*, 359-372.

Lefcourt, H. M., & Martin, R. A. (1986). *Humor and life stress: Antidote to adversity*. New York: Springer-Verlag.

Rust, J., & Goldstein, J. (1989). Humor in marital adjustment. *Humor: International Journal of Humor Research, 2*, 217-223.

Smeltzer, L. R. & Leap, T. L. (1988). An analysis of individual reactions to potentially offensive jokes in work settings. *Human Relations, 41*, 295-304.

Squier, H. A. (1995) Humor in the doctor-patient relationship. *Family Systems Medicine, 13*,101-107.

Thorson, J. A., & Powell, F. C. (1993a) Development and validation of a multidimensional sense of humor scale. *Journal of Clinical Psychology, 49*, 13-23.

Thorson J. A., & Powell, F. C. (1993b). Sense of humor and dimensions of personality. *Journal of Clinical Psychology, 49*, 799-809.

Thorson, J. A., & Powell, F. C. (1994). Depression and sense of humor. *Psychological Reports, 75*, 1473-1474.

Thorson, J. A., Powell, F. C., Sarmany-Schuller, I., & Hampes, W.P. (1997). Psychological health and sense of humor. *Journal of Clinical Psychology, 53*, 605-619.

Valliant, G. E. (1977). *Adaptation to life*. Boston: Little, Brown.

van Wormer, K., & Boes, M. (1997). Humor in the emergency room: A social work perspective. *Health & Social Work, 22*(2), 217-223

Yoder, M. A., Haude, R. A. (1995). Sense of humor and longevity. *Psychological Reports, 76*, 945-946.

Ziv, A. (1984). *Personality and sense of humor*. New York: Springer.

APPENDIX 1. Multidimensional Sense of Humor Scale Scores for Black and White Women.

Item	Whites (N = 200)		Blacks (N = 93)		t
	Mean	SD	Mean	SD	
1. Sometimes I think up jokes and funny stories.	2.2	1.1	2.7	1.1	3.25**
2. Uses of wit or humor help me master difficult situations.	2.6	0.9	2.8	1.0	1.77
3. I'm confident that I can make other people laugh.	2.5	1.0	2.8	0.9	2.66*
4. I dislike comics.	3.4	0.9	3.2	1.1	1.35
5. Other people tell me that I say funny things.	2.8	0.8	2.9	0.9	0.94
6. I can use wit to help adapt to many situations.	2.5	0.9	2.8	0.9	2.42
7. I can ease a tense situation by saying something funny.	2.5	0.9	2.6	0.9	0.80
8. People who tell jokes are a pain in the neck.	3.3	0.8	3.1	1.0	1.29
9. I can often crack people up with the things I say.	2.7	0.9	2.8	0.9	0.26
10. I like a good joke.	3.7	0.5	3.5	0.8	2.09
11. Calling somebody a comedian is a real insult.	3.3	0.8	3.3	0.8	0.18
12. I can say things in such a way as to make people laugh.	2.8	0.8	2.7	0.9	0.81
13. Humor is a lousy coping mechanism.	3.3	0.8	3.3	0.9	0.13
14. I appreciate those who generate humor.	3.5	1.0	3.2	0.7	3.37**
15. People look to me to say funny things.	2.1	1.0	2.4	0.9	2.51
16. Humor helps me cope.	2.9	0.9	2.8	0.9	0.23
17. I'm uncomfortable when everyone is cracking jokes.	2.9	1.3	2.9	1.1	0.19
18. I'm regarded as something of a wit by my friends.	2.4	1.0	2.4	0.9	0.12
19. Coping by using humor is an elegant way of adapting.	2.6	0.9	2.6	0.9	0.33
20. Trying to master situations through uses of humor is dumb.	3.2	0.8	3.0	1.0	1.20
21. I can actually have some control over a group by my uses of humor.	1.9	1.0	2.3	0.9	2.98
22. Uses of humor help to put me at ease.	3.3	0.7	2.8	0.9	5.31**
23. I use humor to entertain my friends.	2.4	1.0	2.7	1.0	2.50
24. My clever sayings amuse others.	2.5	0.9	2.6	0.9	1.26
Total Scale	67.2	12.5	68.2	12.5	0.64

$^*p < .01$ $^{**}p < .001$

APPENDIX 2. Multidimensional Sense of Humor Scale Scores for Black and White Men.

Item	Whites (N=157)		Blacks (N=23)		*t*
	Mean	SD	Mean	SD	
1. Sometimes I think up jokes and funny stories.	2.6	1.0	2.7	1.0	1.66
2. Uses of wit or humor help me master difficult situations.	2.5	0.9	2.7	0.9	0.62
3. I'm confident that I can make other people laugh.	2.8	1.0	2.8	0.9	0.99
4. I dislike comics.	3.2	1.1	3.0	1.2	0.75
5. Other people tell me that I say funny things.	2.9	0.9	2.7	0.8	0.82
6. I can use wit to help adapt to many situations.	2.7	0.9	2.7	0.7	0.03
7. I can ease a tense situation by saying something funny.	2.9	0.9	2.5	0.9	2.15
8. People who tell jokes are a pain in the neck.	3.0	1.0	3.1	1.0	0.22
9. I can often crack people up with the things I say.	3.0	0.9	2.7	0.8	1.66
10. I like a good joke.	3.6	0.9	3.3	1.0	1.14
11. Calling somebody a comedian is a real insult.	3.0	1.0	3.2	0.7	0.91
12. I can say things in such a way as to make people laugh.	3.0	0.9	2.7	0.8	1.39
13. Humor is a lousy coping mechanism.	3.0	1.0	3.0	0.9	0.11
14. I appreciate those who generate humor.	3.2	0.9	2.7	1.0	2.44
15. People look to me to say funny things.	2.3	0.9	2.4	0.9	0.18
16. Humor helps me cope.	2.7	1.1	2.9	0.7	0.76
17. I'm uncomfortable when everyone is cracking jokes.	2.8	1.2	2.8	1.0	0.03
18. I'm regarded as something of a wit by my friends.	2.5	0.9	2.4	0.6	0.49
19. Coping by using humor is an elegant way of adapting.	2.4	0.9	2.4	0.9	0.05
20. Trying to master situations through uses of humor is dumb.	2.8	1.0	2.5	0.9	1.13
21. I can actually have some control over a group by my uses of humor.	2.3	0.9	2.4	1.0	0.69
22. Uses of humor help to put me at ease.	2.3	0.9	2.6	0.9	2.68*
23. I use humor to entertain my friends.	2.7	0.9	2.6	0.8	0.49
24. My clever sayings amuse others.	2.7	0.9	2.7	0.8	0.19
Total Scale	67.7	14.3	65.3	10.6	0.17

* $p < .01$ ** $p < .001$

BEST COPY AVAILABLE

A Comparative Study of Multi-component Weight Loss Treatments: A Replication and Exploratory Look at Female Smokers, Nonsmokers, and Participant BMIs

David L. Johnson
Gary D. Brinker
Southwest Missouri State University

Adult female clients lost highly significant amounts of weight and reduced Body Mass Index (BMI) scores. Treatment I participants who dieted with hypnosis only lost weight and reduced BMI, but lost significantly less weight and BMI than participants treated with overt aversion followed by diet and hypnosis (II). The total sample, I, and II findings strongly replicated earlier studies. Participants were offered I, II, and a third (III) program of overt aversion, then hypnosis, diet, nutritional analyses, and vitamin/mineral supplements. Participants in II and III lost significantly more weight and reduced BMI more than participants in I. Nonsmokers in II and III lost significantly more weight and reduced BMI more than nonsmokers in I. Nonsmokers selected II more often, lost weight and reduced BMI more than smokers in II. Smokers selected I more often than II and III, but lost more weight on average than nonsmokers in III. Obese participants lost more weight and reduced BMI significantly more than overweight participants in the total sample, II, and III.

This field study investigated the respective merits of three, multi-component, weight loss treatment programs. First, the study proposed to replicate the clinical studies, weight-loss, and Body Mass Index (BMI) findings reported earlier (Johnson, 1997; Johnson & Karkut, 1996). Will clients in a multi-component (diet, aversive imagery, hypnosis, behavioral management, counseling, and audiocassette support) weight loss program with an emphasis on overt aversion, then diet and hypnosis, lose more weight and reduce BMI scores more than clients in a multi-component program focused on maintaining a low-calorie diet with hypnosis? Second, the study explored a treatment option with nutritional consultation, laboratory results, and supplements as part of a weight

---

Authors' info: Correspondence should be sent to: David L. Johnson,
1054 E. Seminole Street, Springfield, MO 65807-3040 email:
davej@ipa.net

North American Journal of Psychology, 2001, Vol. 3, No. 1, 13-30.

© NAJP

control program not available in earlier studies. Will clients who use overt aversion, hypnosis, a customized diet and supplements lose weight and reduce BMI scores more than participants in two treatment programs without supplements? In a comparative study (Johnson, 1997) of smoking and nonsmoking participants, adult females in a hypnosis-only treatment achieved highly significant pre- to post-weight losses and reductions in BMI. No statistically significant differences were found between smokers and nonsmokers. Because of the research and program development interests in health, tobacco cessation, and nutrition surrounding smoking and body fat, we looked at possible differences in two subgroups. We compared the use and relative effectiveness of three treatment programs on samples of smokers and BMI classified (overweight and obese) clients. Will overweight and obese nonsmokers lose more weight than overweight and obese smokers? Will obese participants lose weight and reduce BMI scores more than overweight participants by the end of each program?

METHOD

Participants

Over 400 adult female participants were solicited by newspaper advertisements for weight-control services and scheduled for consultations at a private community mental health clinic. After consultation, a total of 388 participants currently not involved in any weight-control, medication, or vitamin/mineral supplement program were scheduled for treatment. Each client selected and paid a clinic fee to participate in one weight-control program. Participants started in either a hypnosis only (*n*=140), an overt aversion and hypnosis (*n*=140) or a combined overt aversion, hypnosis, and nutritional supplement program (*n*=108). The total sample's mean age was 40.3 yr., the mean height was 64.8 in. (1.7 m) and mean starting weight was 169.7 lb. (77.0 kg). The starting mean BMI, computed by taking body weight in kilograms divided by height in meters squared, yielded a total sample mean of 28.4 or approximately 28% overweight (Rowland, 1989).

Measure

Participants completed questionnaires detailing weight history, current height and weight, problem foods and drinks, diet, smoking and health information. Only participants in the Overt Aversion and Hypnosis (OAH) and Combined Overt Aversion, Hypnosis, and Supplements (COAHS) treatments provided examples of tastes, smells, textures, colors, and animals they viewed as disgusting. Participants were weighed without shoes on the same scale on the first day of treatment and weekly thereafter before treatment and at a closing review session.

Procedure

The clinician discussed a low-fat, 1200-1500 calories a day diet approved by the client and client's physician with the participants in the Hypnosis Only (HO) treatment on the second hypnosis session during the first week of the program. The diet was discussed with participants in the OAH treatment program at the first hypnosis session during the second week of the program. The clinician reviewed the nutritional analysis and diet prior to the first hypnosis session after a series of four overt aversion sessions had been completed in the first week of COAHS treatment.

All participants were administered hypnosis treatments in a quiet clinic room while resting in a comfortable recliner. All participants received the same hypnosis content and format used in the three treatment programs. All participants received a self-hypnosis routine before the second hypnosis treatment. Each participant received a copy of the routine to practice outside the clinic before meals and when thinking about or in the presence of foods and drinks. All hypnosis and aversion treatments were administered by the same clinician. All clinic participants received the same 90-minute audiocassette of inductions, visualization routines, suggestions, and weight management segments covered in the program. Participants were instructed to listen to the tape and practice the cognitive routines learned in each program before their scheduled review session. The procedures, diet and treatments used were identical to those reported earlier (c.f., Johnson & Karkut, 1996) with the exception of the COAHS treatment's nutritional component.

Treatment Program I - Hypnosis Only. Participants were hypnotized individually using a standard relaxation induction (e.g., Johnson & Karkut, 1994). During each hypnosis treatment participants were given suggestions referencing the negative consequences of eating high-fat or problematic foods (e.g., candy, desserts, french fries, pork ribs) not on the diet or excess quantities of acceptable foods and drinks. Suggestions were given to support personal attractiveness, health, physical exercise, self-esteem, confidence and commitment to self-regulation, stress reduction, and cognitive-behavioral management of weight goals. Participants in this treatment were scheduled to receive two hypnosis sessions the first week. The diet and self-hypnosis routines were discussed before and suggested during the second hypnosis session. Weekly hypnosis sessions (*M*= 30 min.) were scheduled for the following three weeks with the audiocassette given after the fifth treatment. An exit session with a hypnosis option was scheduled approximately 10 days after the fifth treatment for review of progress, trouble-shooting, use of the audiocassette, and support. The mean duration of treatment from pre- to post-measures was 31.8 days.

Treatment Program II - Overt Aversion and Hypnosis. Participants in this treatment were instructed to bring to the clinic the day before each scheduled aversion session during the first week at least five categories of highly desirable, problem foods (e.g., hot dogs, ice cream, chocolate cake, jellies, syrups, doughnuts) and drinks (e.g., soda, beer, milk shakes). Participants were asked to provide at least five or more cups of each problem category so the clinician could prepare the samples for the next day's scheduled aversion treatment. Participants were told the food and drink samples would be prepared, presented, tasted, and smelled, but not eaten, in an aversive procedure using an electric shock device.

For each aversion session the participant sat in a chair facing a table on which was placed a paper plate, containers of sample foods and drinks, a plastic spoon, and an electric shock device. All samples used in treatment had been prepared by coating them with a mixture of white vinegar and mouthwash to mask the item's taste and smell and to enhance the overt aversiveness of the treatment.

Participants were instructed to stare at the plate while the clinician described the use and safety of a small electric shock device with an adjustable control and two metal electrodes embedded in a Velcro cuff. Electrode gel was applied to the participant's dominant eating hand at the wrist. The cuff was attached securely. Starting at a milliamp reading of 0.0, a level was established within the range of .3 to .9 when a painless, but annoying, involuntary hand movement was elicited by the clinician. Participants were told to place a spoonful of food or drink on the tongue, then close their mouth and eyes. While moving and tasting, but not swallowing the sample, participants were encouraged to shout mentally, 'Make it horrible. Make it disgusting. See it turn into fat.' Each category of food and drink presented for treatment was used for at least three trials during each session. Each participant was shocked from 7 to 11 times at random during each trial. Each category of food and drink was paired with disgusting imagery (e.g., roaches, rat droppings, snakes, spiders) and sickening experiences (e.g., gagging, vomiting) provided by the participant from their aversion questionnaire. These aversive images were used as part of the treatment. Participants were directed to verbalize anger, frustration, and disgust at starchy, greasy, sugary, and artificially colored foods and drinks before purging them in a garbage pail setting next to the table. The mean duration of treatment was 30.1 days.

Treatment Program III - Combined Overt Aversion, Hypnosis, and Supplements. Each participant in this treatment permitted the clinician to take a sample of hair equivalent to one tablespoon along with current dietary information at the initial interview. Participants received the same overt aversion treatment as OAH participants during the first week of the program and weekly hypnosis sessions thereafter. The COAHS participant's sample and dietary information were analyzed independently by a laboratory and a certified nutritionist who prepared a written report for each client. The clinician reviewed the recommendations in the report with the client prior to the first hypnosis session after the scheduled series of overt aversion sessions had been completed. Each individualized report focused on dietary areas with potential health risks and deficiencies (e.g., insufficient calcium, magnesium, iron, zinc, chromium; too much sodium; more vitamins B, C, E). A list of foods and drinks to avoid and specific others to include in the diet were part of the report. Recommendations for vitamin/mineral enriched foods, drinks, and supplements to support nutritional health and stable eating patterns were offered. Participants reported on the weekly use of supplements during the program. At the review session they were encouraged to follow their diet guidelines with periodic medical and nutritional reassessments. The mean duration of treatment was 31.8 days.

RESULTS

A summary of the data used for statistical analyses including *t* ratios for correlated samples on pre- and post-weights and pre- and post-BMI scores for the total sample and by treatment are provided in Table 1. An alpha of .01 was used for all statistical tests. The total sample's pre- to post-weight mean difference of 10.7 lb.(4.9 kg) was highly significant (t_{387} = 40.1, $p < .001$). Each treatment program's pre- to post-weight scores: Treatment Group I mean difference of 9.4 lb. (4.3 kg); Treatment

TABLE 1 Summary of Means and *t* Ratios for Weight (lb) and BMI Scores for Total Sample and by Treatments

	Total Sample *N* = 388	Group I *n*= 140	Group II *n* = 140	Group III *n*= 108
Weight				
Pretest Mean	169.7	163.7	173.6	172.5
Posttest Mean	159.1	154.3	162.3	161.0
Difference	10.7	9.4	11.3	11.5
t^a	40.1**	21.4**	28.8**	20.8**
Body Mass Index				
Pretest	28.4	27.4	29.2	28.8
Posttest	26.6	25.8	27.3	26.8
Difference	1.8	1.6	1.9	2.0
t^a	39.3**	20.6**	27.3**	60.6**

t^a = *t* ratio for correlated samples
**$p < .001$

Group II mean difference of 11.3 lb. (5.1 kg); Treatment Group III mean difference of 11.5 lb. (5.2 kg) were all highly significant ($p< .001$). All but six of the 388 participants lost weight, ranging from 1 (.5 kg) to 32 pounds (14.5 kg.), during the course ($M = 31.4$ days) of treatment. After weight-loss treatment, the entire sample scored a consistent, average drop in BMI ($M = 1.8$). The decline in BMI was highly significant statistically ($p< .001$) across all three treatments and paralleled the pre- to post-weight findings.

TABLE 2 ANOVA of Number of Pounds Lost by Treatments

	N	Mean	*SD*	Std. Error
Treatment I (HO)	140	9.4	5.2	.44
Treatment II (OAH)	140	11.3	4.7	.39
Treatment III (COAHS)	108	11.5	5.7	.55
Total	388.	10.7	5.3	.27

	Sum of Squares	df	Mean Square	F
Between Groups	346.556	2	173.278	6.5*
Within Groups	10318.526	385	28.801	
Total	10665.082	387		

	Multiple Comparisons[1] Mean Difference	Std. Error
Treatment I & Treatment II	- 1.9*	.62
& Treatment III	- 2.1*	.66
Treatment II & Treatment III	- .2	.66

[1] Bonferroni used for all multiple comparisons in this study.
*p < .01

A one-way analysis of variance was used to test for differences in the independent variables of age, height, starting weight, starting BMI, number of treatment sessions, and number of days in a weight-loss program. There were no significant baseline differences among the three treatment groups on age, height, starting weight, number of treatment sessions, or days in a weight loss program. The mean differences between starting BMI scores were significant ($F_{2,385} = 4.3$, $p < .01$). Treatment groups I ($M = 27.4$) and II ($M = 29.2$) yielded the greatest difference in starting BMI.

The dependent variables of total number of pounds lost and reduction in BMI scores were calculated by subtracting respectively the post-weight from the pre-weight scores and the post-BMI from the pre-BMI scores. The ANOVA results, displayed in Table 2, for the three treatment groups on the dependent variable number of pounds lost were statistically significant ($F_{2,385} = 6.5$, $p < .01$). The Treatment I mean of 9.4 lb. (4.3 kg) was significantly different ($p < .01$) from the Treatment II mean (11.3 lb./5.1 kg) and Treatment III mean (11.5 lb./5.2 kg). The Treatment II and III mean difference was not significant.

TABLE 3 ANOVA of Reduction in BMI by Treatments

	N	Mean BMI	*SD*	Std. Error
Treatment I (HO)	140	1.6	.90	.08
Treatment II (OAH)	140	1.9	.80	.07
Treatment III (COAHS)	108	1.9	.98	.09
Total	388	1.8	.90	.05

	Sum of Squares	df	Mean Square	F
Betweeen	9.98	2	4.99	6.3**
Within Groups	303.44	385	.79	
Total	313.42			

	Multiple Comparisons	
	Mean Difference	Std. Error
Treatment I & Treatment II	- .323*	.11
& Treatment III	- .346*	.11
Treatment II & Treatment III	- .023	.11

**p < .01

Because of the significant differences between treatment groups on starting BMI, analyses of covariance were performed on each dependent variable, that is, on pounds lost and on ending BMI, to statistically control for possible effects of this initial participant variable on treatment outcomes. The covariate and dependent variables analyzed met the assumptions of ANCOVA regarding reliability of covariates, linearity, homogeneity, normality, sample size, and outliers (Tabachnick & Fidell, 1989). The analysis of covariance of pounds lost by treatment with starting BMI as covariate was statistically insignificant. The results for the adjusted mean differences among the treatment groups on pounds lost followed the same pattern of results reported for the ANOVA (Table 2). The analysis of covariance on the reduction of BMI scores by treatment with starting BMI as covariate was statistically insignificant. The adjusted mean differences among the treatment groups followed the same pattern of results for BMI reported in Table 3.

Table 4 summarizes the results for t-test comparisons between smokers and nonsmokers in the total sample and for each treatment on the dependent variables pounds lost and reduction in BMI scores over the course of treatment. Nonsmoking participants (*n* = 251) lost an average of .6 lb. (.3 kg) more than a smaller group (*n*=137) of smoking participants. The difference was not statistically significant. There was an insignificant difference in mean pounds (2.0 lb./.9 kg) lost between smokers and nonsmokers in Treatment II favoring nonsmokers. This finding suggests that some of the mean difference in the total sample may have been due to Treatment II's greater impact on nonsmoking subjects' weight loss than smokers in this treatment. Examination of the differences between smokers and nonsmokers in pounds lost by treatments compared to the total sample results by treatments (TABLE 2.) yielded a pattern similar for nonsmokers, but not for smokers. Smokers presented a small mean difference between Treatments I and II, but a larger difference between Treatments II and III. Participants who smoked lost a greater mean number of pounds in Treatment III than did nonsmoking participants.

TABLE 4 Summary of Means and *t* Ratios for Pounds Lost and BMI for Total Sample, Smoking Status, and Treatments

	Total Sample $N = 388$		Group I $n = 140$		Group II $n = 140$		Group III $n = 108$	
	n	Mean	n	Mean	n	Mean	n	Mean
Pounds Lost								
Smokers	137	10.3	66	9.8	40	9.9	31	12.0
Nonsmokers	251	10.9	74	9.1	100	11.9	77	11.3
Difference		.6		.7		2.0		.7
t^b		ns		ns		ns		ns
Mean BMI Reduction								
Smokers	137	1.7	66	1.6	40	1.7	31	2.0
Nonsmokers	251	1.8	74	1.5	100	2.0	77	1.9
Difference		.1		.1		.3		.1
t^b		ns		ns		ns		ns

t^b = *t* ratio for independent samples

The results for the mean reduction in BMI across the three treatment groups yielded a pattern for nonsmokers similar to the original ANOVA (Table 3), but not for smokers. The mean difference (.3) in reduction in

BMI between smokers and nonsmokers in Treatment II was insignificant. Smokers did not, on average, reduce their BMI as much as nonsmokers in Treatment II by program's end.

When participants' treatment choices were cross-tabulated by smoking status, a highly significant ($X^2 = 13.4$, $df = 2$, $p < .001$) distribution of choices resulted. Nonsmokers chose Treatment I (29.5 %), Treatment II (39.8 %), Treatment III (30.7 %), but smokers chose Treatment I (48.2 %), Treatment II (29.2%), Treatment III (22.6%). Clearly a higher percentage of smokers chose Treatment I over Treatments II and III. Treatment II was the choice preferred by nonsmokers.

Analysis of variance of pounds lost by treatment program for nonsmokers (Table 5) was similar to the pattern of results reported in for the total sample. Nonsmokers who participated in Treatment II had significantly ($p < .01$) higher mean weight losses (11.9 lb./5.5 kg) than nonsmokers in Treatment I (9.1 lb./4.1 kg). Nonsmokers in Treatment III had higher mean weight losses (11.3 lb./5.1 kg) than nonsmokers in Treatment I, but the mean difference was not statistically significant. There were no significant differences for nonsmokers between Treatments II and III.

TABLE 5 ANOVA of Total Number of Pounds Lost by Treatments for Smokers and Nonsmokers

	SMOKERS			NONSMOKERS		
	N	Mean	Std. Error	*N*	Mean	Std. Error
Treatment I	66	9.8	.70	74	9.1	.55
Treatment II	40	9.9	.72	100	11.9	.46
Treatment III	31	12.0	.88	77	11.3	.69
Total	137	10.3	.45	251	10.9	.33
	$F = 2.0$			$F = 6.4$*		

Multiple Comparisons

	Mean Difference	Std. Error	Mean Difference	Std. Error
Treatment I & II	- .1	1.04	- 2.8*	.79
Treatment I & III	- 2.2	1.14	- 2.2	.84
Treatment II & III	2.0	1.25	.6	.78

*p < .01

Analysis of variance of pounds lost by treatment program for smokers displayed some differences from nonsmokers. Comparing the sample mean differences for smokers among the treatment groups (Table 5) yielded variations, but no statistically significant findings. Smokers displayed a slightly larger mean difference in pounds lost (2.2 lb./1 kg) between Treatments I and III than between Treatments II and III (2 lb./.9 kg). The smokers' slender mean difference of .1 lb. (.1 kg) between Treatments I and II did not match the nonsmokers' statistically significant (*p*< .01) mean difference of 2.8 lb.(1.3 kg) between Treatments I and II. Smokers achieved a larger mean difference between Treatments I and III. The greatest difference between smokers and nonsmokers occurred in Treatment II. For nonsmokers, as well as the total sample, the difference in total pounds lost between Treatments I and II was greater than the difference between Treatments II and III.

TABLE 6 Summary of Means and *t* Ratios for Pounds Lost and BMI for Total Sample, BMI Status, and Treatments

	Total Sample N = 388		Group I n = 140		Group II n = 140		Group III n = 108	
	n	Mean	*n*	Mean	*n*	Mean	*n*	Mean
Pounds Lost								
BMI < 30	252	9.6	99	8.8	81	10.1	72	10.1
BMI ≥ 30	136	12.5	41	10.7	59	12.9	36	14.3
Difference		2.9		1.9		2.8		4.2
t^b		- 5.5**		- 1.7		- 3.5**		- 3.4**
Mean BMI Reduction								
BMI < 30	252	1.6	99	1.5	81	1.7	72	1.7
BMI ≥ 30	136	2.1	41	1.8	59	2.2	36	2.4
Difference		.5		.3		.5		.7
t^b		- 5.3**		- 1.9		- 3.9**		-3.1*

t^b = *t* ratio for independent samples
*p < .01
**p < .001

We used the published clinical guidelines on the identification of overweight and obese individuals (NIH, 1998) to look for differences between our overweight and obese clients. In order to investigate possible treatment effects among the three programs, we split the sample into two subgroups using the National Institute of Health and World Health Organizations' (1998) recommended BMI cut-off point of 30. Participants with initial BMI less than 30 (low to high overweight) were separated from a comparison group of participants with initial BMI equal to or greater than 30 (low to higher levels of obesity).

Table 6 summarizes the *t*-test results for the overweight and obese participants. For the total sample and each treatment group, obese participants lost more pounds and had higher reductions in BMI scores than overweight participants. For the total sample a highly significant ($t_{386} = - 5.5$, $p < .001$) number of pounds lost favored obese participants by a mean difference of 2.9 lb. (1.3 kg). The mean difference in pounds lost between overweight and obese participants 1.9 lb. (.8 kg) and the reduction in BMI, mean difference .3, in Treatment I were the only comparisons not statistically significant. The mean difference of 2.8 lb. (1.3 kg) in Treatment II was similar to the total sample and highly significant ($p < .001$). Mean differences in BMI reduction for the total sample and Treatments II and III were significant. A similar pattern of results favored obese rather than overweight participants in the reduction of BMI at program's end.

An ANOVA of the number of pounds lost by overweight and obese participants by treatments was statistically insignificant. The small number ($n = 136$) and uneven distribution of obese participants limited interpretation except to report a pattern similar to Table 6. For overweight subjects, the mean differences between treatments were relatively small and statistically insignificant.

DISCUSSION

The total sample of adult female participants displayed strong, consistent, and highly significant ($p < .001$) losses from starting to ending weights and BMIs. Irrespective of treatments, smoking status, and BMI classification, the total sample averaged 10.7 lb. (4.9 kg) lost and a mean BMI reduction of 1.8 points at program's end. Participants had mean weight losses between 7.6 lb. (3.5 kg) for overweight smokers in Treatment II to 14.8 lb. (6.7 kg) for nonsmokers in Treatment III. All subgroups by treatments had statistically significant ($p < .01$) decreases from pre-weight to post-weight using paired *t*-tests. This study replicated the procedures, treatments, and supported findings reported earlier (Johnson & Karkut, 1996; Johnson, 1997) with two multi-component, weight-loss treatment programs. One treatment program focused on introducing a low-calorie diet in the first week of the program with only hypnosis as the main cognitive-behavioral treatment used. The second treatment program focused first on working with problem foods and drinks in a series of four, intense, aversive treatments followed in the second week with an introduction to a low-calorie, reduced-fat diet and weekly hypnosis sessions. The total sample lost 2.4 lb. (1.1 kg) a week, which compares favorably with earlier studies and other short-term, weight loss reports. The total sample's average weight-loss exceeded the suggested standard of 1.2 lb. (.5 kg) a week suggested by Brownell & Wadden (1986) and others.

Hypnosis only participants lost significant amounts of weight and had BMI reductions from start to finish. Even though Treatment I participants started diets on the second session of the first week, 7-10 days earlier than Treatment II and III participants, the reduction in mean weight and BMI scores were less than Treatment II and III participants.

Generally, the data suggest that Treatment II and III consistently outperformed Treatment I participants because they shared the core components of overt aversion followed by hypnosis. In an earlier study (Johnson & Karkut, 1994) of smoking cessation, overt aversion followed by hypnosis treatment was used effectively on the same treatment day. The reverse order, that is, hypnosis first followed by overt aversion, or an overt aversion treatment option only, has not been studied by the investigators. However, these versatile treatment components have been adapted to different clinical, research, and self-instructional purposes (c.f. Johnson & Johnson, 2000).

Treatment III, in addition to the overt aversion and hypnosis components, introduced participants to dietetic consultation, laboratory analysis, and specific nutritional targets in order to look beyond daily calorie counting and a singular focus on weight loss. The perspective was on weight loss and underlying health, nutrition, exercise, and management issues in support of weight maintenance. A nutritionally-sound diet and individualized goals were addressed with recommendations to use specific foods, drinks, and vitamin/mineral supplements. Treatment III yielded a slight, insignificant edge in pounds lost and concurrent reduction in BMI over Treatment II participants. The Treatment III statistical results did not suggest strong, independent variable effects on the dependent measures in the short term.

Obese participants lost more pounds and reduced BMI scores more than nonobese participants in all treatments. Obese participants lost more pounds and reduced BMI scores more in Treatments II and III than in Treatment I. But, obese participants lost more weight on average (*M* = 10.7 lb./4.9 kg) than nonobese participants in Treatment I. The mean reduction in BMI favored obese over non-obese participants in Treatment III and respectively smaller differences in BMI appeared to favor obese over nonobese participants in Treatments II and I. For overweight participants, there were no statistically significant treatment differences.

A review of the total sample's distribution of participants by treatment, smoking status, and BMI classification, yielded a larger number (*n* = 252) of overweight than obese participants in all treatments and a smaller number (*n* = 136) and uneven distribution of obese participants by treatments. A greater number (*n* = 25) and percentage (52.8 %) of obese clients who smoked participated in Treatment I. A much larger number (*n* = 72) and percentage (81.8%) of obese nonsmokers preferred Treatments II and III which limited comparisons of overweight and obese participants by smoking status.

The study replicates and provides empirical support for the consistent clinical weight-loss results by treatment and initial BMI classification. Our prediction would be the greater the starting weight, the more participants will lose in pounds and BMI. Obese participants will lose more than overweight participants across treatments. The findings suggest that the more cognitive-behavioral management and nutritional components available in an effective weight-loss program, the more appeal to a broader range of overweight participants. If it can be shown that these components enhance weight loss, that is, add to the number of pounds lost, reduce BMI scores, and improve nutrition, then these effective components placed in the most workable order for treatment should result in a greater probability of high participant motivation, participation, and rate of weight loss. The more participants perceive program components as connected to their weight-loss issues and relevant to their weight goals, the greater the motivation and participation from individuals at overweight and early to medium obesity levels of starting weight. But what about smokers?

Comparisons between smokers and nonsmokers on pounds lost and reduction in BMI yielded the following. The total sample (n = 251) of nonsmoking participants lost on average (*M* = .6 lb./.3 kg) more pounds and had a reduced difference in BMI scores (*M* = .1) greater than participants who smoked (*n* = 137). In Treatment II nonsmokers lost more pounds (mean difference = 2.0 lb./1 kg) and reduced BMI scores (*M* = .3) more than smokers. Smokers tended to lose pounds and reduce BMI more than nonsmokers in Treatments I and III. Nonsmokers lost on average more pounds in Treatments II and III than Treatment I. Smokers lost more pounds on average in Treatment III, but participated more in Treatment I over II and III. The relatively equal performance in weight loss and reduction in BMI in the Hypnosis Only treatment replicated an earlier study's (Johnson, 1997) findings for smokers and nonsmokers. We were unable to find studies of female smokers participating in weight-loss programs for comparative purposes. In this study smokers lost weight with each treatment program, but the unequal distribution of obese smoking participants across treatments precluded any general statement.

Why the relatively higher rate of weight loss and uneven participation of smokers in Treatments I and III compared to Treatment II participants and nonsmokers? Female participants who smoke may perceive smoking as more important to the management of moods, thoughts (Waters, 2000) appetite, energy, body weight, or self-image than the elimination of problem foods/drinks and their management with aversion, informed food/drink choices, and supplements. From this study we would predict a lower percentage of female smokers would choose a weight loss program which emphasized only the latter components. If participants who smoke viewed the overt aversion treatment component to be too intense, given the already stress-inducing or aversive properties (Parrott, 1999) of smoking, they may have opted first for the stress-reducing, hypnotic, relaxation, and potential sleep benefits in Treatment I. The Treatment I option was presented with less explicit emphasis on aversion and the identification of problem foods/drinks.

Since many smokers eat high-caloric meat, fat consumables, less fruits and vegetables with antioxidants, beta-carotene, vitamins C and E (Ma, 2000) without gaining weight (Perkins, 1992), smokers who begin to gain weight may have been sufficiently concerned to seek treatment. Upon entering treatment, they may have not realized the smoking and fat connections with food, drink, and behavior management (Mikhailidis, 1998). They may have wanted to hold on to the smoking habit, hoping "it" would continue to provide a metabolic lift via release of serotonin (Wurtman & Wurtman, 1995) to manage stress along with fattening foods and drinks (e.g., alcohol, caffeinated beverages) devoid of nutritional value. Many of the participants who smoked may have opted for Treatment I rather than use the overt aversion component.

The overt aversion sessions used in Treatments II and III were acknowledged verbally as disgusting by participants during each session. Participants were sickened by responses to problem foods and drinks as well as when aversive responses were repeated. Also, clients were reinforced by the clinician who paired aversive responses with foods and drinks provided and targeted by participants. Aversive responses were repeated with the aid of each participant's list of problem foods, disgusting items, and sickening experiences. Regardless of smoking status, no participant in the overt aversion sessions asked for more sessions; nor, did any participant request an extra overt aversion session at review.

After participants completed a stressful, aversion component, participants who smoked in Treatment III may have perceived less threat and reduction in stress. They may have gained confidence, then accepted supplements as rewards for their participation, health, nutrition, and psychological well being. All participants in Treatment III may have shared in these perceived benefits, but smokers coping with weight and smoking issues may have received an extra boost in the short-term from these qualitative benefits. Since regular smoking depletes many nutritional reserves and exhausts daily vitamin requirements, smoking participants typically had more mineral and vitamin deficiencies. They benefited from supplements like antioxidant vitamins B, C, E and the amelioration of some toxic effects of daily smoking. Vitamins C and E have been reported (Miller, 1997; Gamble, 2000) to have positive effects on the cardiovascular, respiratory, and circulatory systems of individuals who smoke, even while smoking. A complete B-complex may have helped stressed out participants who smoked relax, stabilize mood, and improve receptivity for subsequent hypnosis and self-hypnosis components. Supplements may have positively affected participants' weight loss in Treatment III by reducing mineral imbalances. Corrected mineral imbalances for overweight smokers included: (1) An improved ratio of sodium to potassium which allowed a natural release of retained water and reduction of sodium. More potassium would support the coordination and efficient use of large muscles in daily activities and exercise with less cramping and fatigue. (2) More chromium may have helped stabilize participants with fluctuating blood sugar levels, hypoglycemia, mood (see “tired arousal” in Gold, *et al.*, 1995), appetite swings, and other precursors of Type II diabetes mellitus (National Institutes of Health, 1998).

In the short run, Treatment III benefits may have helped smokers cope with nutritional, fitness, and health issues associated with smoking, lower rates of weight lost, and BMI reduction. Smokers who chose Treatment I did not receive the overt aversive component, supplements, and prospective treatment benefits. Overweight and obese smokers in Treatment III lost on average 2-3 lb. more than overweight and obese smokers in Treatment I.

When we focus on weight-loss participants who regularly smoke, the results by treatment and initial BMI classification are not consistently supported and remain generally untested. Nonsmokers lost on average more weight and had greater reductions in BMI when compared with participants who smoked in all treatments for obesity and especially in Treatment II. A slight reversal of results occurred in favor of overweight (BMI < 30.0) smokers losing slightly more weight on average than overweight nonsmoking participants in Treatments I and III. Questions on how to control for self-reported smoking rates and smokers' participation in weight loss and BMI studies have not been resolved (c.f., Calle, 1999). We believe the unexpected weight losses by smokers warrant further systematic study. How the complexities of metabolism, weight gains and losses interface with the daily exigencies of persons caught up in a smoking cycle who want to lose weight before, or instead of, smoking cessation deserve further systematic study and innovative program applications.

The overweight BMI classification itself might have accounted for some participation by smokers in treatment, for it is the class in which symptoms first appear in association with health risks, co-morbid, and lipid-related conditions. Depression, glucose intolerance, diabetes mellitus II, hypertension, hyperlipidaemia, and other cardiovascular indicators can often be associated with the overweight classification. We strongly recommend that clinicians and health-care providers share information on BMI, health risks, and associated co-morbid conditions that emerge with unwanted weight gains and associated tobacco smoking. We want to encourage and support tobacco cessation before anyone starts a weight loss program. In the absence of tobacco cessation, we hope clinicians and prospective clients would consider effective weight-loss treatments modeled on one or more of the three effective treatments outlined herein. We have some optimism that clinical research, modest weight loss, BMI reductions, nutritional support, and weight-loss maintenance can work towards clients' tobacco cessation, enhanced health, fitness, and positive lifestyle changes.

Obese participants in the study may have experienced many chronic metabolic signs including more past problems with specific foods and drinks which, in part, may have strengthened and focused their motivation for treatment, especially Treatments II and III. If overweight participants and overweight smokers in the study became acutely aware of multiple risk factors, it may have accounted for stronger motivation to enter and complete weight-loss treatment. Participants in our study were not randomly assigned to treatment so more systematic, controlled studies are needed to test these hypotheses. The literature strongly suggests that the underlying health and fitness of weight-loss participants including smokers at every BMI level, plus lifestyle conditions outside the clinic walls, must fill in significant pieces to the weight-loss puzzle. The twin, major preventable epidemics of obesity and tobacco addiction continue to thrive in our expanding sedentary lifestyle, population, and global environment.

REFERENCES

Brownell, K. D., & Wadden, T. A. (1986). Behavior therapy for obesity: Modern approaches and better results. In K. D. Brownell & J. P. Foreyt (Eds.), *Handbook of eating disorders: physiology, psychology, and treatment of obesity, anorexia, and bulimia.* (pp. 180-197) New York: Basic Books.

Calle, E. E., Thun, M. J., Petrelli, J. M., Rodriguez, C., & Heath, Jr., C. W. (1999). Body-mass index and mortality in a prospective cohort of U.S. adults. *The New England Journal of Medicine, 341*, 1097-1105.

Gamble, J., Grewal, P. S., & Gartside, I. B. (1998). Vitamin C modifies the cardiovascular and microvascular responses to cigarette smoke inhalation in man. *Clinical Science, 4*, 455-460.

Gold, A. E., MacLeod, K. M., Frier, B. M., & Deary, I. J. (1995). Changes in mood during acute hypoglycemia in healthy participants. *Journal of Personality and Social Psychology, 68*, 498-504.

Johnson, D. L., & Karkut, R. T. (1994). Performance by gender in a stop-smoking program combining hypnosis and aversion. *Psychological Reports, 75*, 851-857.

Johnson, D. L., & Karkut, R. T. (1996). Participation in multi-component hypnosis treatment programs for women's weight loss with and without overt aversion. *Psychological Reports, 79*, 659-668.

Johnson, D. L. (1997). Weight loss for women: Studies of smokers and nonsmokers using hypnosis and multi-component treatments with and without overt aversion. *Psychological Reports, 80*, 931-933.

Johnson, D. L., & C. A. Johnson (2000). *Stop smoking and chewing tobacco for life changes*. Bryn Mawr, PA: BUYBOOKSontheweb.com.

Ma, J., Hampl, J. S., & Betts, N. M. (1994-1996). Antioxidant intakes and smoking status: Data from the continuing survey of food intakes by individuals. *American Journal of Clinical Nutrition, 71*, 774-780.

Miller, E. R. III, Appel, L. J., Jiang, L., & Risby, T. H. (1997). Association between cigarette smoking and lipid peroxidation in a controlled feeding study. *Circulation, 96*, 1097-1101.

Mikhailidis, D. P., Papadakis, J. A., & Ganotakis, E. S. (1998). Smoking diabetes and hyperlipidaemia. *Journal of Research and Social Health, 118*, 91-93.

National Institutes of Health (1998). *Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults: The Evidence Report*, NIH Publication No. 98-4083.

National Task Force on the Prevention and Treatment of Obesity (2000). *Archives of Internal Medicine, 160*, 898-904.

Parrott, A. C. (1999). Does cigarette smoking cause stress? *American Psychologist, 54*, 817-820.

Perkins, K. A. (1992). Effects of tobacco smoking on caloric intake. *British Journal of Addictions, 87*, 193-205.

Rowland, M. L. (1989). A nomogram for computing Body Mass Index. *Dietetic Currents, 16*, 1-12.

Tabachnick, B. & Fidell, L. (1989). *Using Multivariate Statistics*. New York: HarperCollins Publishers, Inc.

Waters, A. J., & Sutton, S. R. (2000). Direct and indirect effects of nicotine/smoking on cognition in humans. *Addictive Behavior*, *25*, 29-43.

Wurtman, R. J., & Wurtman, J. J. (1995). Brain sertonin, carbohydrate-craving, obesity, and depression. *Obesity Research*, *4*, 477S-480S.

World Health Organization (1998). Obesity: Preventing and managing the global epidemic. In *Report of a World Health Organization Consultation on Obesity*, Geneva, 3-5 June, 1997, Geneva, Switzerland: World Health Organization, 1-276.

Convergent Validity of the Situational Outlook Questionnaire: Discriminating Levels of Perceived Support for Creativity

Scott G. Isaksen and Kenneth J. Lauer
Creative Problem Solving Group - Buffalo

The purpose of this study was to continue examining the validity of the Situational Outlook Questionnaire (SOQ™). The relationship between mean ratings of the nine dimensions that make up the Situational Outlook Questionnaire and the self-perception of the organizational climate as conducive (or not conducive) to creativity was tested with a sample of 1,830 individuals from a variety of organizations. Participants were categorized on their level of perceived support for creativity. The results indicated that the means of all nine dimensions of the Situational Outlook Questionnaire were different for each level of perceived support for creativity in the work environment. The difference was statistically significant for each level. The results of this study indicated that the Situational Outlook Questionnaire may be able to discriminate effectively among different levels of perceived support for creativity in the immediate work environment.

The purpose of this study was to examine the relationship between an individual's perceptions of the degree of support for their personal creativity and the climate for creativity and change where they work. The study utilized a modified version of the Situational Outlook Questionnaire (SOQ™) to conduct the inquiry. The SOQ is a paper and pencil instrument intended for use as a diagnostic tool to improve awareness and understanding of the organization's ability to support creativity and change. It is conceptually grounded in previous empirical and theoretical work on creativity and innovation (Drucker, 1985; Ekvall, 1996; Lauer, 1994).

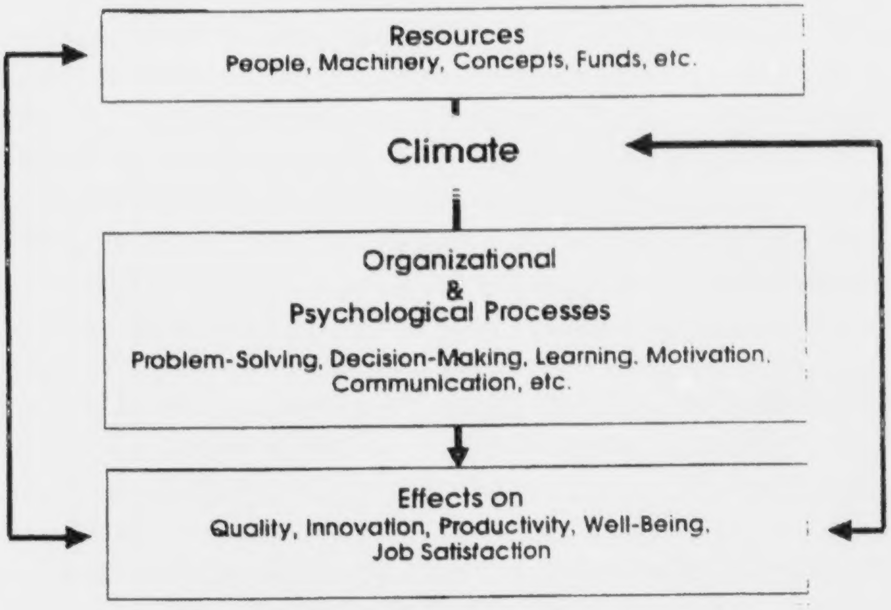
The basic theoretical model underlying the SOQ proposes that creative productivity, in any context, is dependent upon numerous individual, group, and organizational variables. These variables combine to influence the patterns of behavior in any work environment. These recurring patterns, as well as their underlying attitudes and feelings, characterize the atmosphere or "quality of life" within an organization and is defined as organizational climate (Ekvall, 1996; Isaksen, Lauer, & Ekvall, 1999).

Authors' info: Correspondence should be sent to: S.G. Isaksen c/o Creative Problem Solving Group-Buffalo; 1325 No. Forest Rd., Suite 340, Williamsville, NY 14221. cpsb@cpsb.com

North American Journal of Psychology, 2001, Vol. 3, No. 1, 31-40.

© NAJP

Ekvall (1991) asserted that climate acts as an intervening variable in an organization. Climate influences, and is subsequently influenced by, the outcome of organizational operations. Climate affects outcomes by influencing organizational processes such as problem solving, decision making, communicating and coordinating, the individual processes of learning and creating, and levels of motivation and commitment. These in turn influence the ways in which the organization uses its resources, such as people, buildings, intellectual property and funds. These effects subsequently become apparent in the quality of the products or services produced, whether these are radically new products, minor improvements of old ones, job satisfaction, productivity, profit, or emotional and physical well-being. These factors, in turn, affect both the availability of resources and the climate itself (See Figure 1).


Resources
People, Machinery, Concepts, Funds, etc.
Climate
Organizational & Psychological Processes
Problem-Solving, Decision-Making, Learning, Motivation, Communication, etc.
Effects on
Quality, Innovation, Productivity, Well-Being, Job Satisfaction

Adapted from Ekvall, Swedish Council for Management and Work

FIGURE 1 A Model of Climate as an Intervening Variable

The environment may be either conducive or detrimental to creativity. Stein (1968) expressed this orientation by stating, "creativity, like all behavior, is a function of the transactional relationships between the individual and his environment" (p.936). The climate within organizations that stimulates creative productivity has been the subject of increasing attention and research (Amabile & Gryskiewicz, 1989).

This study was designed to examine the nature of the relationship between individuals' perceptions of support for their personal creativity and their results on the nine dimension scores of the SOQ. Our specific research interest was in determining whether a significant relationship existed between the individual's perception of the supportiveness of their own immediate work environment for creativity and each of the nine dimensions of the SOQ.

METHOD

Participants

The study used data from 1,830 individuals; of these 1,469 were members of organizations that contracted for training courses in Creative Problem Solving (CPS) with the Creative Problem-Solving Group - Buffalo (CPS-B), while 361 were participants in educational and organizational research samples.

Regarding sex, 54.8% were male, 45.2% were female. The mean age of those respondents (*n* = 977) who reported their age was 35.3 years with a range of 17 to 64 years. The educational level of the sample, reported by 762 subjects, was spread over the following groups: completed high school (*n*=21), some college education (*n* = 134), bachelor's degree (*n* = 249), some graduate education (*n*=22), master's degree (*n* = 177), and doctorate (*n* = 159).

Participants came from a wide range of organizational levels and functions. The majority of participants included in this study came from six organizations (*n* = 1,469), while the remaining subjects were undergraduate students who had previously attended courses on creativity (*n* = 195), attended educational conferences (*n* = 122), or attended five-day creative problem-solving courses (*n* = 44). Participants represented a variety of organizations, namely: 74%' from Business/Goods Industries (*n* = 1354), 3% from Education (*n* = 55), .6% from the Government or Service sector (n = 9), 10.7% were college students (*n* = 196), and 11.8% did not specify their type of organization (*n* = 216).

The majority of the participants in the Business/Goods Industry category came from two large international organizations. Several courses were conducted at a petroleum company, accounting for a total of 755 respondents, or 41.3% of the total sample. Programs conducted for a consumer products manufacturer provided another 528 subjects, or 28.9 % of the total sample.

Measures

The SOQ is a revised translation of the Creative Climate Questionnaire (CCQ; Isaksen, Lauer, Murdock, Dorval, & Puccio, 1995). The former is designed to assess nine aspects of organizational climate that either foster or hinder creative behavior and organizational change.

The initial translation of the questionnaire from Swedish to English began in 1986. Two independent Swedish translators then used a process known as back translation with decentering to adjust the initial translation (Bontempo, 1993). The English translation was then reviewed for face validity using a Q-sort by the initial translation team, and six creativity researchers, one from India, two from Norway, and three from the USA.

The questionnaire that resulted from this translation process was similar to Ekvall's CCQ in that it contained five items for each of the 10 theoretical dimensions. The senior author then used this as a research instrument. Data collected from 419 individuals who completed this first 50-item version of the SOQ were analyzed by Lauer (1994). An exploratory principal component analysis (Varimax rotation) revealed 10 factors with an eigenvalue equal to or greater than 1.0 and accounted for 62.1% of the total variance. Examination of the delineation of items in the Varimax rotation did show some inconsistency with the theoretical loading patterns. The coefficient alphas for the sample ranged from .72 to .87 for the 10 theoretically based factors. A study by Isaksen and Kaufmann (1990) reported similar coefficient alphas for a sample of 634 individuals.

Following this initial translation and analysis process, the SOQ was tested and refined on four different versions of the measure. This process is described in more detail in the technical manual (Isaksen, Lauer, Murdock, Dorval, & Puccio, 1995) and Cabra's thesis (1996). The intent of this refinement process was to improve the factor structure and coefficient alphas. Exploratory factor analysis supported a nine-factor principal axis (oblique rotation) factor structure rather than the 10 factor principal component (varimax rotation) structure reported by Ekvall, Arvonen, and Waldenstrom-Lindblad (1983). This nine factor structure is reported by Isaksen, Lauer, & Ekvall (1999) in a study of 1,111 individuals who used the fourth version of the SOQ. Coefficient alphas for the nine factors ranged from .62 to .89. Further studies of the SOQ's reliability and validity have been conducted (Britz, 1995; Grivas, 1996; Isaksen & Kaufmann, 1990; Isaksen et. al., 1995; Talbot, Cooper, & Barrows, 1992; Turnipseed, 1994).

The SOQ is intended for adult respondents. Flesch Reading Ease scores for the questionnaire instructions and the items were 54.4 and 59.3 respectively. This converts to a USA grade-school reading level of 8.3 for the instructions and 7.6 for the items.

In 1996 the SOQ was made available for use with groups and organizations. However, because of the multidimensional nature and intended use of the questionnaire, it must be administered and debriefed by individuals who are qualified and trained to use the theory and measure for effective interventions.

The SOQ version used in this study consisted of positive dimensions (Challenge, Freedom, Idea Support, Playfulness/Humor, Debate, Trust/Openness, Risk-taking, and Idea Time) which tend to foster a more creative climate. The Conflict dimension, as defined, is considered to be detrimental to creativity in the workplace. It is important to distinguish Conflict from Debate. Conflict focuses on negative emotional and personal tensions, whereas Debate centers upon differences of ideas, viewpoints, and individually unique experiences and knowledge. Both dimensions incorporate tension; debate pertains to idea tension, while conflict entails personal tension.

The version used in this study contained five items for each of the nine SOQ dimensions and the omnibus question cited later in this section. Of the 46 items, 22 items are reverse scored to control for response bias. The items are framed in such a manner that they ask the respondent to be an objective observer of the environment in which he/she is working. Respondents answer the items on a 4-point scale in which 0 = Not at all applicable, 1 = Applicable to some extent, 2 = Fairly applicable, 3 = Applicable to a high degree. The overall scores for each dimension are calculated by taking the mean of the participant's response for each dimension and multiplying this by 100. All dimensions therefore, have a theoretical range from 0 to 300. This procedure allows for ease of comparison across dimensions. The coefficient alphas for the SOQ's nine theory-based factors obtained from this sample were as follows: Challenge, .81; Freedom, .69; Conflict, .72; Idea Support, .83; Playfulness/Humor, .78; Debate, .82; Trust/Openness, .71; Risk-taking, .52; and Idea Time, .81.

Studies of validity have used demographic grouping and/or self-report responses as a means of classifying respondents into meaningful groupings (Litwin & Stringer, 1968; Sackett & Larson, 1990; Taylor & Gryskiewicz, 1993). In this study an omnibus question was used to classify respondents into meaningful groups. The question was phrased: "I feel the immediate work environment is supportive of my personal creativity." The respondents were divided into four groups depending on their responses to this question. The group of participants who had scored their environment as "not supportive" was labeled Not Supportive (n = 201). Those in the second group found their immediate work environment conducive to their creativity "to some extent" and are labeled To Some Extent (n = 609). The third group consisted of respondents who had answered the omnibus question with "fairly applicable" and are labeled Fairly Supportive (n = 702). Finally, those in fourth group said that the description was "applicable to a high degree" to their work environment and are labeled Highly Supportive (n = 318).

Procedure

All participants completed the questionnaire individually focusing their responses on their perception of their immediate working environment. The SOQ was distributed through the various organizations mail systems along with a memorandum describing the purpose of the measure. The participants from educational and organizational research studies were given the questionnaire by the senior author in classroom settings and the questionnaires were collected immediately upon completion. Completed instruments were then returned to the CPS-B for scoring and analysis. Steps were taken to ensure voluntary participation and confidentiality.

TABLE 1 SOQ Means and SDs Sorted by Support of the Immediate Work Environment to Personal Creativity With Results of One-way ANOVAs Between Groups*

Dimension	Not Supportive (n=201)		To Some Extent (n=609)		Fairly Supportive (n=702)		Highly Supportive (n=318)			
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	F**	R^2
Challenge	149.2	68.5	187.3	56.4	224.3	50.0	257.7	45.8	221.33	.27
Freedom	135.9	51.6	168.8	47.4	204.6	42.8	237.7	45.4	275.14	.31
Idea Support	109.8	57.0	167.8	50.2	213.0	49.5	251.0	49.9	406.10	.40
Playfulness / Humor	103.1	54.8	143.9	50.9	179.9	50.0	216.3	55.5	252.52	.29
Debate	127.7	61.2	167.2	54.1	201.1	49.1	233.1	48.0	216.80	.26
Trust / Openness	98.0	47.8	140.6	45.4	176.3	45.0	215.3	49.3	334.94	.36
Conflict	178.3	63.7	136.2	58.8	108.2	53.6	77.4	53.4	158.53	.21
Risk-Taking	112.0	48.2	139.8	40.4	166.8	35.9	194.2	44.2	223.68	.27
Idea Time	77.7	55.8	115.1	50.0	150.6	51.9	193.5	57.6	256.62	.30

*df = 3, 1826

** all p values < .001

RESULTS

The means and standard deviations of the four groups on each dimension, as presented in Table 1, indicate that the scores on all the SOQ dimensions increased as the individuals perceived their immediate work environment as more supportive to their personal creativity (except for the Conflict dimension, which decreased). To determine if differences were statistically significant between the four groups, we conducted a one-way multivariate analysis of variance (MANOVA). This was followed by nine one-way analyses of variance (ANOVA's) for the nine dimensions to determine the strength and statistical significance of the differences among all four categories on each dimension. The MANOVA results were statistically significant [Wilks's $\Lambda = .46$, $F(3,1826) = 53.54$, $p < .001$]. The subsequent one-way ANOVA's (See Table 1) showed that this effect was statistically significant for all dimensions. The *F*-value for each dimension was statistically significant at $p < .001$. The magnitude of the effect sizes as measured via R^2 were all greater than .26 and may be considered large (Cohen, 1977).

DISCUSSION

The preceding results provide evidence that a statistically significant difference does exist between the respondents' reports of perceived support for creativity and the scores for the SOQ dimensions. The results show that individuals who perceive their immediate work environment as more supportive of their creativity report higher mean scores on those dimensions of the SOQ that are positively related to the creative climate. Lower mean scores were observed for the Conflict dimension for those who report more support for their personal creativity. This suggests that there is a meaningful relationship between the manner in which individuals perceive their work environment and how they view the environment's ability to support their personal creativity in the organizational settings included in this study.

There are, however, several limitations to this study. First, only a single item was used to categorize the subject groups. Additional omnibus items may be added to obtain an improved sense of how the respondents perceive the overall creative climate in their immediate work environment. It may also be useful to examine further how the SOQ dimensions relate to other measures such as job satisfaction and other measures of organizational climate, creative climate, and culture. For example, Turnipseed (1994) found that several job satisfaction factors and social environmental variables are related to the SOQ dimensions.

The coefficient alpha of two dimensions is below .70, which is commonly used as an acceptable baseline for a research instrument (Freedom, .69, and Risk-Taking, .52). The SOQ version used in this study was the third revision. Subsequent revisions of the questionnaire have addressed this issue and improvements have been reported in the literature (Isaksen, et al., 1999). Future research will continue exploring means to improve the questionnaire's reliability.

Another limitation to this study is that the ratings of the overall environmental support for creativity were measured by self-reports. Additional research using outside observers and a variety of alternative measures are needed to further validate the dimensions of creative climate assessed by the SOQ. Further utility can be realized from a study of this nature if a large random sample were acquired and analyzed, since the findings and insights could be transferred to other organizational settings with higher levels of confidence. Finally, since climate is conceived as an attribute of the organization, future validation studies should also focus on organizational performance measures.

An additional limitation is that both the criterion measure (omnibus question) and the SOQ were combined and completed by respondents at the same time. Shared method variance may result and this may be an issue in a study of this nature. Further research should take this into account.

Despite these limitations this study has many merits. The large sample size and the wide variety of organizations it contains provides a comparative base for other studies on the topic to build upon. The sample size also allows further studies to be conducted that may address how perceived level of support for creativity relates to the type of industry, the intensity of the technology used, and the rate of change of technology in that industry. These and other questions may also be studied in relation to the departments within an organization or the functions individuals perform in the organization.

In summary, the results of this study coupled with the psychometric findings of other studies (Isaksen & Lauer, 1999; Isaksen, et al., 1999; Isaksen, et al, 1995) support the continued investigation into the SOQ and how it can be used to further our understanding of creativity and change in organizational settings.

REFERENCES

Amabile, T. M., & Gryskiewicz, N. D. (1989). The creative environment scales: Work Environment Inventory. *Creativity Research Journal, 2,* 231-253.

Bontempo, R. (1993). Translation fidelity of psychological scales: An item response theory analysis of an individualism-collectivism scale. *Journal of Cross-Cultural Psychology, 24,* 298-314.

Britz, A. (1995). *The assessment of climate for innovation in organizations.* Unpublished master's thesis, Technische Hochshule Darmstadt, Germany.

Cabra, J. F. (1996). *Examining the reliability and factor structure of the Climate for Innovation Questionnaire.* Unpublished master's thesis, State University College at Buffalo, New York.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences.* San Diego, CA: Academic Press.

Drucker, P. E. (1985). *Innovation and entrepreneurship: Practice and principles.* New York: Harper & Row.

Ekvall, G. (1996). Organizational climate for creativity and innovation. *European Journal of Work and Organizational Psychology, 5,* 105-123.

Ekvall, G. (1991). The organizational culture of idea-management: A creative climate for the management of ideas. In J. Henry, & D. Walker (Eds.), *Managing innovation* (pp. 73-79). London: SAGE Publications Ltd.

Ekvall, G., Arvonen, J., & Waldenstrom-Lindblad, I. (1983). *Creative organizational climate: Construction and validation of a measuring instrument.* (Report 2). Stockholm, Sweden: FArådet - The Swedish Council for Management and Work Life Issues.

Grivas, C. C. (1996). *An exploratory investigation of the relationship of cognitive style with perceptions of creative climate.* Unpublished Masters Thesis. Buffalo, NY: Center for Studies in Creativity.

Isaksen, S. G., & Kaufmann, G. (1990). Adaptors and innovators: Different perceptions of the psychological climate for creativity. *Studia Psychologica, 32,* 129-141.

Isaksen, S. G., Lauer, K. J., & Ekvall, G. (1999). Situational Outlook Questionnaire: A measure of the climate for creativity and change. *Psychological Reports, 85,* 665-674.

Isaksen, S. G., & Lauer, K. J. (1999). Relationship between cognitive style and individual psychological climate: Reflections on a previous study. *Studia Psychologica, 41,* 177-191.

Isaksen, S. G., Lauer, K. J., & Ekvall, G. (1998). *Perceptions of the best and worst climates for creativity: Preliminary validation evidence for the Situational Outlook Questionnaire.* (Monograph #305). Buffalo, NY: Creative Problem Solving Group - Buffalo.

Isaksen, S. G., Lauer, K. J., Murdock, M. C., Dorval, K. B., & Puccio, G. J. (1995). *Situational outlook questionnaire: Understanding the climate for creativity and change (SOQ™) - A technical manual.* Buffalo, NY: Creative Problem Solving Group - Buffalo.

Lauer, K. J. (1994). *The assessment of creative climate: An investigation of Ekvall's Creative Climate Questionnaire.* Unpublished master's thesis, State University College at Buffalo, New York.

Litwin, G. H., & Stringer, R. A. Jr. (1968). *Motivation and organizational climate.* Boston, MA: Harvard University.

Sackett, P. R., & Larson, J. R. Jr. (1990). Research strategies and tactics in industrial and organizational psychology. In M. D. Dunnette, & L. M. Hough E. F. (Eds.), *Handbook of industrial & organizational psychology* (pp. 419-489). Palo Alto, CA: Consulting Psychologists Press, Inc.

Stein, M. I. (1968). Creativity. In E. F. Borgatta, & W. W. Lambert (Eds.), *Handbook of personality theory and research* (pp. 900-942). Chicago, IL: Rand McNally.

Talbot, R., Cooper, C., & Barrows, S. (1992). Creativity and stress. *Creativity and Innovation Management, 1,* 183-193.

Taylor, S., & Gryskiewicz, N. (1993). A test of the validity of the Work Environment Inventory. *Educational and Psychological Measurement, 53,* 557-563.

Turnipseed, D. (1994). The relationship between the social environment of organizations and the climate for innovation and creativity. *Creativity and Innovation Management, 3,* 184 - 195.

Winer, B. J. (1971). *Statistical principles in experimental design*. New York: McGraw-Hill.

Note: This study was conducted in part while Albert Winsemius was a visiting graduate student at the Center for Studies in Creativity and a research intern at the Creative Problem Solving Group - Buffalo. We wish to thank him for his efforts. The authors would also like to thank Dr. Donald Treffinger for his comments on this paper and Barbara Babij, Marves Isaksen, and Wendy Lee Reddy for their editorial suggestions. We would also like to thank the editor and reviewers of the *North American Journal of Psychology* for their helpful comments and suggestions.

Alcohol Consumption and Consequences in a Sample of University Undergraduates: Patterns and Relations to Internalized Distress and Religious Affiliation

J. M. Oliver Cynthia K. S. Reed
Bruce W. Smith
St. Louis University

This study examined self-reports of alcohol use and other psychological problems common in undergraduates in 248 students at a midsize Jesuit university using the "Core Survey," designed specifically for undergraduates. Drinking and Its Problems emerged as a factor independent of both internalized distress and eating dysfunctions. Drinking and Its Problems was decomposed into six factors: Heavy Drinking; High-Risk Behaviors; General Role Impairment; Sexual Victimization Due to Alcohol; Aggression Due to Alcohol; and Failing at Quitting. Catholic religious affiliation, male gender, and on-campus residence significantly predicted Drinking and Its Problems. Results suggest student drinking is attended by problems and related to Catholic religious affiliation and that campus personnel need to know of these relationships.

Although rates of consumption among undergraduate students taken as a group are currently stable (Presley, Meilman, & Lyerla, 1993, 1995), investigation of alcohol use in college students remains important. Evidence regarding many fundamental issues regarding student alcohol use remains inconclusive. Such issues include; whether student drinking, although heavy, is uncomplicated by problems; whether some student drinking may resemble alcohol dependence; and whether Catholic religious affiliation is a risk factor for heavy consumption and/or problems related to drinking.

One of the reasons why evidence about these significant issues remains inconclusive is that investigators of student drinking have tended to develop their own instruments for assessing it, making comparisons across studies difficult. A second reason is that investigators have been slow to apply multivariate techniques to the analysis of responses of what is likely the best instrument available for assessing alcohol use in students, the Core Drug and Alcohol Survey ("Core Survey"; Presley et

---

Author info: Correspondence should be sent to: J. M. Oliver, Department of Psychology, St. Louis University, 208 Shannon Hall, 3544 Laclede Blvd., St. Louis, MO 63103. E- mail to oliverjm@primary.net.

North American Journal of Psychology, 2001, Vol. 3, No. 1, 41-60.

© NAJP

al., 1993, 1995). A third reason is that investigators have shown some reluctance to examine the relation between religious affiliation and alcohol use.

Patterns of use in college students are still little understood. According to The Diagnostic and Statistical Manual-Fourth Edition (DSM-IV; American Psychiatric Association, 1994), the characteristics which differentiate alcohol use (a normal phenomenon) from alcohol abuse (a mental disorder associated with impairment) are a pattern of heavy consumption and recurring problems due to heavy consumption. Following DSM-IV, four general classes of problems differentiate between heavy consumption unaccompanied by problems and alcohol abuse: failure to fulfill obligations; exposing oneself to a physical hazard (high-risk behaviors); incurring legal problems; and incurring social or interpersonal problems.

The distinction between heavy consumption, or "heavy drinking," and problems due to heavy consumption, or alcohol abuse, is important in undergraduates as well as in non-student adolescents and young adults. Although student status is a well-known risk factor for high rates of consumption, drinking in college students, or at least one variant of it, may be limited to "heavy drinking." Students may drink heavily during their undergraduate years for experimental and recreational purposes without incurring problems related to heavy consumption, then decrease their consumption and their risk of developing attendant problems when they leave the campus and the student role. By contrast, drinking in college students, or at least one variant of it, may be associated with problems just as problems are associated with non-student adolescents and young adults who drink heavily.

The handful of studies which have examined patterns of drinking in college students (Brennan, Walfish, & AuBuchon, 1986a, 1986b; Brown, 1985; Park, 1967) have found from two to 25 patterns or factors of use, of which at least one was generally a heavy-drinking pattern uncomplicated by attendant problems. Evidence regarding "heavy drinking" in students uncomplicated by problems has been summarized recently by Engs, Diebold, and Hanson (1996); Wechsler, Dowdall, Davenport, and Castillo, (1995); Wechsler, Davenport, Dowdall, and Moeykens, (1994). Oxford, Waller, and Peto (1974) have also found that heavy drinking in students is not necessarily associated with problems.

By contrast, however, the majority of evidence suggests that undergraduate drinking is accompanied by problems (Abbey, 1991; Anderson & Mathieu, 1996; Buelow & Koeppel, 1995; Caroll & Caroll, 1995; Desiderato & Crawford, 1995; Finley & Corty, 1993; Hughes & Dodder, 1983; Meilman, 1993; Rivinus, 1987; Rivinus & Larimer, 1993;

Wechsler et al., 1995; Wechsler et al., 1994). A problem of particular concern ensuing from alcohol consumption is high-risk behaviors (Buelow & Koeppel, 1995; Rivinus, 1987; Rivinus & Larimer, 1993; Wechsler et al., 1994; Wechsler et al., 1995) high-risk sexual behaviors (Anderson & Mathieu, 1996; Caroll & Caroll, 1995; Desiderato & Crawford, 1995; Meilman, 1993; Wechsler et al., 1995; Wechsler et al., 1994) and sexual assault (Abbey, 1991; Finley & Corty, 1993).

Patterns of alcohol use in undergraduates may include not only alcohol use associated with problems but alcohol dependence. According to DSM-IV, the characteristics that differentiate alcohol abuse from alcohol dependence consist of physiological tolerance, withdrawal symptoms associated with efforts to decrease consumption, and various manifestations of compulsive use. Park (1967), who studied patterns of alcohol use, identified a problem-drinking pattern that might predict future alcohol dependence.

Some demographic variables may make good predictors of maladaptive alcohol use and allow students at risk to be identified early. Demographic correlates of alcohol use in undergraduates have been examined frequently. The single demographic variable most strongly associated with alcohol consumption and with problems associated with alcohol use in college students is gender. A very large body of evidence indicates that college males drink in larger amounts, with greater frequency, and experience problems associated with alcohol use in greater numbers and with higher severity; for a recent summary of this evidence, see Engs et al. (1996).

Religious affiliation is another demographic variable that is related to alcohol consumption and to associated problems. In general, acknowledgment of some religious affiliation irrespective of sect has been found to be negatively associated with alcohol consumption (Cherry, 1987; Cronin, 1995; Engs et al., 1996). However, data regarding Catholic religious affiliation as a risk factor for alcohol use are beginning to be published (Engs, et al., 1996; Hanson, 1977). Catholic religious affiliation has been found to predict both higher levels of consumption and a higher incidence of problems related to drinking (Engs et al., 1996) and has also been found to be associated with earlier age of initial drinking (Hanson, 1977). One unpublished study found significantly higher rates of consumption among undergraduates at Jesuit institutions than among those at non-Jesuit ones (Wernig, 1989).

Thus far, studies have examined the association between Catholic religious affiliation and alcohol consumption in one of two ways: either they have investigated the relation between religious affiliation and alcohol consumption among students at a variety of types of university, both sectarian and non-sectarian; or they have compared rates between Jesuit and non-Jesuit institutions. A third way in which the relation between Catholic religious affiliation and consumption may be investigated is within a Catholic university where a substantial proportion of non-Catholic students are enrolled.

It remains unclear whether alcohol consumption in undergraduates results from other behavioral and psychological problems, such as anxiety, depression, and eating problems (Lundholm, 1989; MacDonald, Fleming, & Barry, 1991; Pullen, 1994; Ross & Tisdall, 1994). Drinking might be secondary to and might mask other primary problems that would need to be identified early in order to reduce secondary drinking or "self-medication." Conversely, drinking might be a prelude to and culminate in psychological problems such as anxiety, depression, and/or eating problems. In the latter case, alcohol consumption would be a primary problem and would need to be identified early in order to prevent secondary problems from accruing. Some investigators, however, have found little or no relation between adjustment and consumption (Fondacara & Heller, 1983; Kashubeck & Mintz, 1996; Kim, Lamier, Walker, & Marlatt, 1997).

An instrument developed specifically for the study of undergraduate substance use is the Core Survey. The Core Survey was developed under the auspices of the Fund for the Improvement of Postsecondary Education (FIPSE) to gather baseline and trend data in college populations regarding a broad range of facets of substance use and abuse in undergraduate students and to allow historical trends to be examined (Presley et al., 1993, 1995). The Core Survey thus examines both alcohol consumption and consequences of alcohol use.

Although the Core Survey has been used to track alcohol consumption and consequences of alcohol use in college students, including through time (Presley et al., 1995), responses to the survey have apparently not yet been examined from a multivariate perspective. The Core Survey thus provides a unique opportunity to examine the association between consumption and consequences; patterns of use; and the possible presence of patterns consistent with alcohol abuse and dependence. Further, responses to the Core Survey have not yet been examined in relation to religious affiliation; indeed the Core Survey has no item inquiring about religious affiliation. Thus a striking opportunity to ascertain the association between religious affiliation and alcohol use in college students has so far been neglected.

An important feature of the Core Survey is that each question regarding a potential consequence of alcohol use asks specifically whether the consequence was experienced as a direct result of alcohol use. The Core Survey therefore allows the investigator to discriminate between problems that might occur for other reasons and problems specific to alcohol use. Another important feature of the Core Survey is that to some degree it covers all four types of problems that the DSM-IV includes as problems critical to alcohol abuse. Although only two questions are directed to role impairment, the areas of high-risk behaviors, legal problems, and social problems are each covered by a minimum of three questions each. An additional six questions are addressed to other types of consequences that are not specifically catalogued in DSM-IV. Another important feature of the Core Survey is that two items inquire about problems that might pertain to alcohol dependence as distinct from alcohol abuse; for example, one question asks whether the respondent has already tried unsuccessfully to stop using alcohol.

Our institution, a midsize, private, Catholic, Jesuit university in the Midwest, affords a good setting in which to examine these issues. We examined levels of alcohol consumption, consequences of alcohol consumption, anxiety, depression, physical symptoms, and eating problems in the same population. This allowed us to investigate the degree to which these psychological problems common in undergraduate populations are risk factors for (as distinct from independent of) alcohol consumption and consequences of alcohol use. About 70% of our undergraduates are Catholic, providing an excellent population in which to study the potential relation between religious affiliation and alcohol use.

METHOD

Participants

Participants were 248 undergraduate volunteers who were recruited in a variety of ways, from phone calls to their residence to solicitation by their resident advisor. Although a portion of the sample was randomly selected, the majority of the sample was not. The participation rate was approximately 33%. The type of recruitment of participants and method of administration (whether group or individual, in a standard lecture-hall or individual residence setting) was recorded.

Table 1 displays demographic characteristics of participants. The typical participant was female, Caucasian, Catholic, and a full-time student. Year levels were almost equally represented.

Participants were compared to population statistics for undergraduates at this university. These comparisons indicated that the sample was representative of the undergraduate population at this institution in most ways, but that males were proportionally underrepresented. Because we wanted our findings to be as generalizable as possible, we applied a

TABLE 1 Demographic Characteristics of the Sample

Variable	Percent
Gender	
Male	26.9
Female	73.1
Age	
16 or younger	0.4
17	0.4
18	17.1
19	29.0
20	16.3
21-24	28.6
25-29	3.3
30-39	2.9
40-54	2.0
Year Level	
Freshman	28.6
Sophomore	23.7
Junior	22.9
Senior	24.9
Ethnicity	
Hispanic	1.2
Asian/Pacific Islander	4.5
White	85.4
Black	6.9
Other	2.0
Religion	
Catholic	62.0
Jewish	0.4
Protestant	15.1
Other	22.4
Residence	
On-Campus	57.8
Off-Campus	42.2

weighting factor to compensate statistically for under-representation of male gender in all subsequent data analyses conducted across gender. This weighting factor provided a final sample which was 44% male and 56% female, characteristic of our own and likely many other institutions. Comparisons were also made among participants recruited and tested in various ways. Despite rather large differences in methods of recruitment and administration, no differences were found among the different groups of participants.

Measures

Core Alcohol and Drug Survey ("Core Survey"): Facets of substance use assessed by the Core Survey (Presley et al., 1993, 1995) include the nature of substances used and abused; the amount of substances and frequency with which they are consumed; and negative consequences of students' substance use. The Core Survey was developed to provide information regarding specific behaviors and beliefs; data are reported by individual question, and subscales are not available.

The Core Survey is a comprehensive self-report instrument that includes but is not limited to a survey of five aspects of consumption of 12 substances. Items pertaining to substance consumption and consequences of consumption have responses in a Likert format with varying numbers of levels depending on the type of item. Correlations between items pertaining to alcohol use and items pertaining to consequences of alcohol consumption ranging from .08 to .65 and test-retest correlations over an unspecified time interval ranging from .59 to 1.00 have been reported.

Beck Anxiety Inventory (BAI): The BAI (Beck, Epstein, Brown, & Steer, 1988) was developed specifically to measure the severity of anxiety and to discriminate anxiety from depression reliably in psychiatric populations. The BAI is a self-report inventory that assesses 21 common symptoms of anxiety using a four-point Likert scale; scores range from 0 to 63. It has excellent internal reliability, as measured by coefficient alpha, and high test-retest reliability. Excellent convergent validity and discriminant validity have been reported in psychiatric populations (Beck et al., 1988; Fydrich, Dowdall, & Chambless, 1992). High internal consistency reliability and comparable factor structures have been demonstrated in both clinical and student samples (Borden, Peterson, & Jackson, 1991).

Beck Depression Inventory (BDI): The BDI (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) was designed to assess severity of affective, cognitive, motivational, and physiological aspects of depressive symptoms in psychiatric populations. The BDI is a 21-item self-report measure whose response alternatives have a Likert format in a four-point continuum; scores range from 0 to 63. The BDI has been validated for use with a university population (Bumberry, Oliver, & McClure, 1978).

Student Adaptation to College Questionnaire (SACQ): The Personal-Emotional Adjustment subscale was selected from the SACQ (Baker &

Siryk, 1989) for use in this study because of its capacity to assess positive adjustment as the potential opposite pole of dysphoria. The SACQ was developed for two purposes: to identify students experiencing adjustment difficulties who might benefit from remedial intervention; and to provide a source of dependent variables in investigations of students' adjustment to college as functions of personality and of environmental variables.

The SACQ is a 67-item self-report measure whose responses are provided on a 9-point Likert scale. For the Personal-Emotional Adjustment subscale, in 25 independent samples of freshpersons at 10 institutions, the mean internal consistency reliability as measured by Cronbach's alpha was .82. Regarding concurrent validity, the Personal-Emotional Adjustment subscale has been found to correlate substantially with other indices of positive adjustment; most particularly, it has been found to correlate .80 with the Mental Health Inventory (Baker & Siryk, 1989).

The Stress Audit: The Stress Audit (Miller, Smith, & Mehler, 1983) is a comprehensive measure of three aspects of stress: 1) external stressors; 2) subjective stress; and 3) relative vulnerability to stress operationalized primarily as health-related behaviors and lifestyle. Five stress scales were selected to measure physical symptoms because they comprise a comprehensive and systematic review of symptoms that may occur in each of five bodily systems. A sixth scale, Emotional Stress, was selected to measure dysphoria that included anxiety and depression. Respondents rated the amount of stress caused by each physical symptom, during the last six months on a five-point Likert scale. In a large sample of patients referred to a stress management clinic, internal consistency reliability coefficients for subscales ranged from .76 to .98. Cronbach's alpha for composite scores ranged from .87 to .93; test-retest reliability coefficients for intervals from one to six weeks ranged from .48 to .92, and concurrent validity ranged from .46 to .70 (Miller, Smith, & Mehler, 1983).

Eating Disorder Inventory (EDI): The EDI (Garner, Olmstead, & Polivy, 1983) was designed to measure behaviors, attitudes, and psychological traits common in anorexia nervosa and bulimia. The EDI was originally designed to assess two related but discriminable aspects of eating dysfunctions, which it measures on two types of subscales: subscales pertaining to behaviors and attitudes; and subscales pertaining to personality traits that characterize people with anorexia nervosa and bulimia. Behavioral and attitudinal subscales consist of Drive for Thinness; Bulimia; and Body Dissatisfaction. Scores on Drive for Thinness, Bulimia, and Body Dissatisfaction were combined into a single composite score ("Eating Problems"). Personality trait subscales consist of Ineffectiveness; Perfectionism; Interpersonal Distrust; Interoceptive

Awareness; and Fear of Maturity. Scores on Ineffectiveness, Perfectionism, Interpersonal Distrust, Interoceptive Awareness, and Fear of Maturity were combined into a single composite score of personality traits associated with eating problems ("Eating Traits").

The EDI is a 64-item self-report measure on which responses are indicated on a six-point Likert scale. Internal consistency reliability of subscales as measured by Cronbach's alpha > .80 and convergent validity ranging from .43 to .68 (p <.001) have been reported (Garner, Olmstead, & Polivy, 1983). In non-clinical populations in general and in college populations in particular, it has been found to have high internal consistency and to be a multidimensional instrument with a relatively stable factor structure (Crowther, Lilly, Crawford, & Shepherd, 1992; Klemchuk, Hutchinson, & Frank, 1990; Raciti & Norcross, 1987).

Cooperative Institutional Research Program's Student Information Form ("CIRP"): The CIRP was used to obtain participants' demographic characteristics. Demographic variables surveyed on the CIRP included gender; age; class standing; parental income; and residence. Religious affiliation was surveyed in a question designed specifically for this population.

Procedure

Measures and their respective instructions were combined into a single questionnaire. Participants recorded their responses on an answer sheet that could be scanned optically. Participants were asked whether they would be willing to participate in a follow-up study. Those who declined to participate in follow-up were promised complete anonymity; those who agreed to participate in follow-up were promised complete confidentiality.

RESULTS

Table 2 lists means and standard deviations on measures of psychological problems by gender, together with results of tests of the significance of differences between genders, and for the sample as a whole.

In order to ascertain the relation between alcohol consumption and consequences of alcohol use, we factor analyzed participants' responses to all measures. We selected principal components analysis with Oblimin rotation because there is no previous evidence about the possible structure of psychological problems or possible relations between alcohol consumption, consequences of alcohol consumption, internalizing problems, and eating problems. Factor analysis was conducted within and across genders because of the possibility that psychological problems might cohere in different patterns as a function of gender.

TABLE 2 Means and Standard Deviations on Measures of Psychological Variables by Gender and Results of Tests of Differences Between Genders

	Males		Females		
Measure	M	SD	M	SD	F
Beck Anxiety Inventory	8.79	8.92	11.49	8.32	7.28**
Beck Depression Inventory	7.60	6.54	8.72	7.17	1.9
Emotional Stress	15.23	12.54	20.97	13.17	14.54**
Personal-Emotional Adjustment	90.52	16.85	86.11	21.36	3.71
Physical Symptoms	44.02	29.39	55.67	33.97	9.66**
Alcohol Consumption	15.57	7.97	12.24	6.02	17.07**
Alcohol Consequences	30.79	11.20	27.57	10.09	6.62*
Eating Problems	6.02	7.16	18.11	15.77	64.37***
Eating Traits	16.77	11.76	18.97	14.46	1.97*
Internalized Distress	12.27	47.96	36.16	69.84	15.86***
Alcohol Use and Its Consequences	42.69	16.27	34.23	12.26	13.67***

*$p < .05$
**$p < .01$
*** $p < .001$

Factors extracted within males, within females, and across genders were very similar; we therefore report results across genders as shown in Table 3. In the total sample, two factors were extracted which together accounted for two-thirds of the variance. Factor I, which accounted for about half of the variance, was characterized by high loadings of Personal-Emotional Adjustment (loading negatively), the BDI, and the BAI; by high moderate loadings of Eating Traits, Emotional Stress, and Physical Symptoms; and by a considerable loading of Eating Problems. Factor I appeared to be characterized best as “Internalized Distress”. Factor II, which accounted for an additional 18% of the variance, was defined by high loadings of Alcohol Consumption and Alcohol Consequences and seemed best termed "Alcohol Use and Its Problems".

Alcohol Use and Its Problems was only minimally correlated with Internalized Distress (*r* =.10, *n.s.*). Alcohol consumption and consequences of alcohol use thus cohered together into a single integrated factor appropriately construed as Alcohol Use and Its Problems, which was almost completely independent of symptoms of anxiety and depression, physical symptoms, and eating-disordered behaviors and attitudes. Eating Problems were not definitively represented by the factors extracted.

TABLE 3 Factor Loadings of Psychological Problems on Factors Extracted from Indices by Gender and for Total Sample

	Males Factor			Females Factor		Total Sample Factor	
	I	II	III	I	II	I	II
Beck Anxiety Inventory	.76			.83		.81	
Beck Depression Inventory	.72			.88		.82	
Emotional Stress	.88			.81		.77	
Personal/ Emotional Adjustment	-.41		.60	-.85		-.83	
Physical Symptoms	.82			.72		.74	
Alcohol Consumption		.94			.92		.93
Alcohol Consequences		.89			.82		.87
Eating Problems			-.92	.54		.56	
Eating Traits			-.78	.82		.78	
Eigenvalue	3.84	1.68	1.23	4.46	1.50	4.17	1.62
Percent of Variance Accounted For	42.6	18.6	13.7	49.5	16.6	46.3	18.0

Participants' scores on Alcohol Use and Its Problems (as well as on Internalized Distress) by gender, together with results of the test of significance of the difference between the genders, and for the total sample are displayed in Table 2. Males scored significantly higher than did females on Alcohol Use and Its Problems.

We next examined alcohol use in its dual aspects of consumption and consequences at a finer level of analysis. To permit this level of analysis, we decomposed the composite index of consumption into the four items that contributed to the consumption index and decomposed the composite index of consequences into the 19 items that contributed to the consequences index. DSM-IV's criteria for Alcohol Abuse embed "recurrent" or "continued use" of alcohol despite the occurrence of negative consequences in each of the four positive or "rule-in" criteria. This precedent provided a rationale for including the four indices of frequency and amount of alcohol consumption in the analysis with the 19 consequences of consumption.

TABLE 4 Factor Loadings of Indices of Alcohol Consumption and Consequences on Factors Extracted From Total Sample

Index of Alcohol Use	Fact. 1	Fact. 2	Fact. 3	Fact. 4	Fact. 5	Fact. 6
Frequency in Past Year	.96					
Frequency in Past 30 Days	.94					
Episodes of Binge Drinking/ Past 30 Days	.79					
No. of Drinks/ Past Week	.84					
Poor Performance, Test Project			-.93			
Missed Class			-.65			
Had Hangover	.55					
Got Nauseated/ Vomited			-.32			
Drove While Intoxicated	.57					
Had Memory Loss				-.42		-.34
Seriously Thought of Suicide				-.72		
Serious Tried to Commit Suicide		.79		-.33		
Was Hurt or Injured			-.49		-.45	
InTrouble W/ Police/Authority		.39				
Damaged Property				.31		
Arrested for DWI/DUI		.78				
Got into Argument or Fight			-.80			
Was Criticized by Someone			-.74			
Was Taken Advantage of Sexually				-.82		
Took Advantage of Another Sexually					.84	
Thought Might Have Drinking Problem						-.72
Did Something Regretted Later			-.62			
Tried Unsuccessfully to Stop						-.89
Eigenvalue	7.60	2.42	2.14	1.36	1.26	1.09
% of Variance Accounted For	33.10	10.54	9.30	5.90	5.50	4.80

We simplified and organized data from the 23 discrete indicators of alcohol use by means of factor analysis. We again selected principle components analysis with Oblimin rotation as an exploratory, non-hypothesis-testing method of factor analysis that could be compared with results obtained by grouping consequences according to the scheme used in DSM-IV. Results of the factor analysis are displayed in Table 4.

As shown in Table 4, six factors emerged from the analysis. Taken together, the six factors accounted for nearly 70% of the variance. Oblique rotation failed to produce simple structure; only the first two factors were defined by items with positive loadings. However, the factors were psychologically meaningful, so results of oblique rotation were preferred to those of a Varimax rotation.

Factor I was clearly the strongest factor, accounting for one-third of the variance. Factor I was defined by the four indices of frequency and amount of consumption, all of which loaded very highly on the factor. In addition, the (short-term, physiological) consequence of having a hangover loaded moderately highly on Factor I, as did another short-term but potentially very serious consequence of driving while intoxicated. Factor I therefore seemed to be described appropriately as "heavy drinking"; it is debatable whether Factor I can appropriately be construed as "heavy drinking without major problems" because driving while intoxicated is a high-risk behavior. Factors II and III were only about one-third as strong as Factor I, each accounting for about an additional 10% of the variance. Factor II was characterized by high loadings of being arrested for driving under the influence of alcohol and by making a serious suicidal attempt and by moderate loadings of getting in trouble with the police or with campus authorities. Factor II might best be captured by the phrase, "High-Risk Behaviors". Factors III though VI were all characterized by negative loadings. Factor III was defined by very high loadings of having performed badly on a test or other important project due to alcohol use and of arguing or fighting due to use; by high moderate loadings of having been criticized by others for one's behavior while under the influence and having missed a class due to alcohol use; and by moderate loadings of regretting one's behavior and having been hurt or injured due to alcohol use. Factor III appeared to represent impairment in academic and social roles due to externalizing behaviors, some of them active, such as getting into arguments and fights, some of them passive, such as performing poorly or missing classes. Factor III seemed to be appropriately construed as "General Role Impairment". Factors IV, V, and VI were likewise equal in strength, each accounting for about half as much variance as either Factor II or Factor III, or about 5% of the variance. Factor IV, which was also characterized by negative loadings, was defined by high loadings of having been taken advantage of sexually and suicidal ideation due to alcohol use and by a moderate loading of memory loss. Factor IV therefore seemed clearly to represent "Sexual Victimization Due to Alcohol". Factor V was defined by high loadings of having taken advantage of another sexually and damaging property, and by a moderate loading of experiencing injury due to alcohol use. Factor V appeared to be the mirror image of Factor IV and to be captured by the term "Aggression Due to Alcohol". Factor VI was defined by high loadings of having tried unsuccessfully to stop using alcohol and thinking that one might have an alcohol problem. Factor VI seemed therefore to represent "Failing at Quitting."

TABLE 5 Summary of Stepwise Multiple Regression Analyses for Variables Predicting Alcohol Abuse Across and Within Genders

Variable	*B*	*SE B*	*Beta*
		Total Sample (N = 288)	
Catholic Religion	9.16	1.692	.305***
Gender	-5.283	1.638	-.180***
Residence	-3.831	1.664	-.131*
		Males (*N* = 61)	
Catholic Religion	14.015	4.228	.396**
		Females (*N* = 172)	
Catholic Religion	5.349	1.872	.214**

Notes: for Total Sample, $R^2 = .09$ for Step 1; change in $R^2 = .03$ for Step 2; change in $R^2 = .02$ for Step 3.
For Males, $R^2 = .16$. For Females, R = .05.
* $p < .05$ ** $p < .01$ **** $p < .0001$

We sought correlates of Alcohol Use and Its Problems as a dimension of psychological problems because we wanted to learn whether Catholic religious affiliation might help to identify students at risk. We selected four other demographic variables as ones that would compete with Catholic religious affiliation to help reveal the strength of religious affiliation as a predictor. We selected gender because virtually all previous research has found male gender to be associated with greater consumption and more serious consequences of use in undergraduates (Engs et al., 1996); ethnicity, because African-American ethnicity has been found to be associated with less alcohol use in college students (Presley et al., 1993, 1995); residence, because on-campus residence has been associated with higher rates of alcohol use; and parental income because, in our sample, zero-order bivariate correlations showed that it was significantly related to Alcohol Use and Its Problems.

In order to make religious affiliation and ethnicity, the two non-ordinal variables with more than two levels, appropriate for use in multiple regression analyses, we dummy-coded them, folding into an "other" category those categories with very few participants. More specifically, we dummy-coded religious affiliation into three levels: Catholic, Protestant, and other. Ethnicity was dummy-coded into into white, African-American, Asian or Pacific Islander, Hispanic, and other. We calculated participants' factor scale scores on Alcohol Use and Alcohol Consequences and summed participants' scores across both measures.

Multiple regression analyses were then conducted across and within genders using Alcohol Use and Its Problems as the criterion and the set of demographic variables as predictors. Results of multiple regression analysis are shown in Table 5. Alcohol Use and Its Problems was significantly predicted by Catholic religion, which predicted about 10% of the variance; gender, which predicted an additional 3% of the variance, and on-campus residence, which predicted an additional 2% of the variance. These three variables taken together explained about 15% of the variance in Alcohol Use and Its Problems in the total sample. Alcohol Use and Its Problems was significantly predicted by a sole predictor, Catholic religious affiliation, within both males and females. In males, Catholic religious affiliation accounted for about 16% of the variance; in females, about 5% of the variance.

We also sought to establish correlates of the six factors of alcohol use by means of multiple regression. We used as predictors the same variables used to predict the composite dimension of Alcohol Use and Its Problems as described above, using each of the six factors of alcohol use as criteria in turn. Results of these multiple regression analyses were remarkably similar to those obtained in seeking correlates of the factor Alcohol Use and Its Problems; religious affiliation and gender repeatedly emerged as predictors.

DISCUSSION

In this sample of undergraduate students, Alcohol Use and Its Problems emerged as a single coherent pattern of psychological and behavioral problems, which was statistically independent of internalized distress. Levels of Alcohol Use and Its Problems were higher in males than in females.

Items contributing to Alcohol Use and Its Problems could be decomposed into six psychologically meaningful patterns of drinking. Of these, the first and strongest was "heavy drinking", consisting of heavy consumption complicated by short-term problems such as having a hangover and driving while intoxicated. It is debatable whether this factor can be considered as relatively uncomplicated heavy drinking, given that driving while intoxicated can have potentially extreme consequences.

The next four factors extracted were all associated with problems attributable to alcohol use, problems that included high-risk behaviors, general role impairment, sexual victimization due to alcohol consumption, and aggression due to alcohol consumption.

The final factor extracted from items contributing to Drinking and Its Problems consisted of inability to stop drinking. The presence of this factor suggests that alcohol use may be difficult to stop for some undergraduates and may progress to alcohol dependence.

Among demographic risk factors for alcohol abuse, the strongest in our sample at a Jesuit institution was Catholic religious affiliation. Catholic religious affiliation exceeded male gender, a previously well established demographic predictor of alcohol consumption and problems associated with alcohol use, as a predictor of Alcohol Use and Its Consequences. On-campus residence also significantly predicted Alcohol Use and Its Consequences.

The present data fail to provide support for either the hypothesis that those who drink heavily are doing so to self-medicate for problems such as anxiety and depression, or the observation that drinking continued over a sufficient period of time results in subjective distress, particularly anxiety and depression.

The emergence of Catholic religious affiliation as a significant predictor of Alcohol Use and Its Problems joins three other reports of Catholic religious affiliation as a risk factor for alcohol consumption and for problems associated with alcohol use (Engs et al., 1996; Hanson, 1974; Wernig, 1989). These four investigations have now examined the relation between Catholic religious affiliation and alcohol use in students enrolled at various types of institutions, both non-sectarian and sectarian; compared consumption rates in students enrolled at Jesuit and non-Jesuit institutions; and ascertained levels of alcohol use and consequences of alcohol use among students of varying religious affiliations enrolled at a single Jesuit institution. All three types of study have found Catholic religious affiliation to be a significant correlate of alcohol consumption and/or problems related to alcohol use. It might be noted as well that Catholic religious affiliation predicted alcohol abuse within both men and women and predicted specific patterns of Alcohol Use and Its Problems ranging from heavy drinking through failing at quitting.

No study has yet separated Catholic religious affiliation from ethnicity, which is frequently correlated with Catholic religious affiliation. A study currently underway within the same Jesuit institution will separate Catholic religious affiliation from ethnicity. Likewise, no investigation has so far been reported which tests the strength of the interaction between Catholic religious affiliation and gender as a predictor of alcohol consumption and/or problems associated with alcohol use. Certainly the differential magnitude of the association between Catholic religious affiliation and Alcohol Use, and Its Problems found within gender in our sample, suggests that the interaction between religious affiliation and gender is likely to be significant in this population.

This study is the first to our knowledge to apply multivariate techniques to responses to the very carefully designed Core Survey. It is likewise the first to our knowledge to examine alcohol consumption and consequences of alcohol use in the context of other psychological problems common in undergraduate students. Finally, it is the first to our knowledge to relate religious affiliation to alcohol consumption and consequences of use as assessed by the Core Survey.

Limitations of this study include the reliance on self-report measures of alcohol consumption and consequences of alcohol use and the relatively small sample of males. Although the self-report measure of alcohol use selected for this investigation has been used extensively and was carefully selected, further investigations of alcohol use in undergraduates might make use of collateral informants such as roommates and friends. We compensated statistically for the relatively small number of males who participated in this study, but it would be desirable to obtain a larger absolute number of male participants.

Findings from this investigation imply that because of the virtually complete independence of internalized distress and Alcohol Abuse and Its Problems, students who are drinking to excess and/or experiencing problems related to drinking are unlikely to be motivated to present themselves for treatment, whether at a counseling center sponsored by the institution or elsewhere. Rather, patterns of alcohol use and alcohol-related problems identified in this investigation suggest that students who are consuming alcohol heavily are likely to have academic, social, disciplinary, and legal problems, problems that will come to the attention of faculty, resident advisors, and administrators. Although resident advisors and administrators may already be aware of associations between these problems and heavy alcohol consumption, faculty may be less aware of them. It may be helpful to educate faculty about such connections, about ways of addressing such issues with students who have problems, and about appropriate referrals when the statistical association between academic problems and maladaptive alcohol use seems likely to apply in the case of a particular student. Additionally, because of the growing evidence that Catholic religious affiliation predicts alcohol consumption and problems associated with alcohol use at a variety of different types of educational institutions, it may be both feasible and necessary to develop psycho-educational programs regarding risks of drinking behavior targeted specifically for Catholic students.

REFERENCES

Abbey, A. (1991). Acquaintance rape and alcohol consumption on college campuses: How are they linked? *Journal of American College Health, 39,* 165-169.

American Psychiatric Association. (1994). *Diagnostic and Statistic Manual-Fourth Edition.* Washington, D.C.: American Psychiatric Association.

Anderson, P. B., & Mathieu, D. A. (1996). College students' high-risk sexual behavior following alcohol consumption. *Journal of Sex and Marital Therapy, 22,* 259-264.

Baker, R. W., & Siryk, B. (1989). *Student Adaptation to College Questionnaire Manual.* Los Angeles, CA: Western Psychological Services.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology, 56,* 893-897.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561-571.

Borden, J. W., Peterson, D. R., & Jackson, E. A. (1991). The Beck Anxiety Inventory in nonclinical samples: Initial psychometric properties. *Journal of Psychopathology and Behavioral Assessment, 13,* 345-356.

Brennan, A. F., Walfish, S., & AuBuchon, P. (1986a). Alcohol use and abuse in college students: I A review of individual and personality correlates. *International Journal of the Addictions, 21,* 449-474.

Brennan, A. F., Walfish, S., & AuBuchon, P. (1986b). Alcohol use and abuse in college students: II. Social/environmental correlates, methodological issues, and implications for intervention. *International Journal of the Addictions, 21,* 475-493.

Brown, S. A. (1985). Expectancies versus background in the prediction of college drinking patterns. *Journal of Consulting and Clinical Psychology, 53,* 123-130.

Buelow, G., & Koeppel, J. (1995). Psychological consequences of alcohol induced blackout among college students. *Journal of Alcohol and Drug Education, 40,* 10-20.

Bumberry, W., Oliver, J. M., & McClure, J. N. (1978). Validation of the Beck Depression Inventory in a university population using psychiatric estimate as the criterion. *Journal of Clinical and Consulting Psychology, 46,* 150-155.

Caroll, J. L., & Caroll, L. M. (1995). Alcohol use and risky sex among college students. *Psychological Reports, 76,* 723-726

Cherry, A. L. (1987). Social bond theory and alcohol use among college students. *Journal of College Student Personnel, 28,* 128-135

Cooperative Institutional Research Program [CIRP]. (1992). *National Norms Report.* Los Angeles, CA: University of California at Los Angeles Higher Education Research Institute, 1992.

Cooperative Institutional Research Program [CIRP]. (1995). *National Norms Report.* CA: University of California at Los Angeles Higher Education Research Institute, 1995.

Cronin, C. (1995). Religiosity, religious affiliation, and alcohol and drug use among American college students living in Germany. *International Journal of the Addictions, 30*, 231-238

Crowther, J. H., Lilly, R. S., Crawford, P. A., & Shepherd, K. L. (1992). The stability of the Eating Disorder Inventory. *International Journal of Eating Disorders, 12*, 97-101.

Desiderato, L. L., & Crawford, H. J. (1995). Risky sexual behavior in college students: Relations between number of sexual partners, disclosure of previous risky behavior, and alcohol use. *Journal of Youth and Adolescence,24*, 55-68.

Engs, R. C., Diebold, B. A., & Hanson, D. J. (1996). The drinking patterns and problems of a national sample of college students, 1994. *Journal of Alcohol and Drug Education, 41*, 13-33

Finley, C., & Corty, E. (1993). Rape on campus: The prevalence of sexual assault while enrolled in college. *Journal of College Student Development, 34*, 113-117.

Fondacaro, M. R., & Heller, K. (1983). Social support factor and drinking among college student males. *Journal of Young and Adolescence, 12*, 285-299.

Fydrich, T., Dowdall, D., & Chambless, D. L. (1992). Reliability and validity of the Beck Anxiety Inventory. *Journal of Anxiety Disorders, 6*, 55-61.

Garner, D. M., Olmstead, M. P., & Polivy, J. (1983). Developments and validation of a multidimensional eating disorder inventory for anorexia nervosa and bulimia. *International Journal of Eating Disorders, 2*, 15-34.

Hanson, D. J. (1977). Trends in drinking attitudes and behaviors among college students. *Journal of Alcohol and Drug Education,* 22, 17-22

Hughes, S. P., & Dodder, R. A. (1983). Alcohol-related problems and collegiate drinking patterns. *Journal of Youth and Adolescence, 12*, 65-76.

Kashubeck, S., & Mintz, L. B. (1996). Eating disorder symptomatology and substance use in college females. *Journal of College Student Development, 37*, 396-404

Kim, E. L, Larimer, M. E., Walker, D. D., & Marlatt, G. A. (1997). Relationship of alcohol use to other health behaviors among college students. *Psychology of Addictive Behaviors, 11*, 166-173.

Klemchuk, H. P., Hutchinson, C. B., & Frank, R. T. (1990). Body dissatisfaction and eating-related problems on the college campus: Usefulness of the Eating Disorder Inventory with a nonclinical population. *Journal of Counseling Psychology, 37*, 297-305.

Lundholm, J. K. (1989). Alcohol use among university females: Relationship to eating disordered behavior. *Addictive Behaviors, 14*, 181-185.

MacDonald, R., Fleming, M. F., & Barry, K. L. (1991). Risk factors associated with alcohol abuse in college students. *American Journal of Drug and Alcohol Abuse, 17*, 439-449.

Meilman, P. W. (1993). Alcohol-induced sexual behavior on campus. *Journal of American College Health, 42*, 27-31.

Miller, L. H., Smith, A. D., & Mehler, B. L. (1983). *The Stress Audit Manual.* Brookline, MA: Biobehavioral Associates, 1415 Beacon Street, 02146.

Oxford, J., Waller, S., & Peto, J. (1974). Drinking behavior and attitudes and their correlates among university students in England: I. Principal components in the drinking domain. II. Personality and social influences. III. Sex differences. *Quarterly Journal of Studies on Alcohol, 35*, 1316-1374.

Park, P. (1967). Dimensions of drinking among male college students. *Social Problems, 14*, 473-482.

Presley, C. A., Meilman, P. W., & Lyerla, R. (1993). *Alcohol and Drugs on American College Campuses: Use, Consequences, and Perceptions of the Campus Environment. Vol. I: 1988-1990.* Carbondale, IL: The Core Institute.

Presley, C. A., Meilman, P. W., & Lyerla, R. (1995). *Alcohol and Drugs on American College Campuses: Use, Consequences, and Perceptions of the Campus Environment. Vol. II: 1990-1992.* Carbondale, IL: The Core Institute.

Pullen, L. M. (1994). The relationships among alcohol abuse in college students and selected psychological/demographic variables. *Journal of Alcohol and Drug Education, 40*, 316-350.

Raciti, M. C., & Norcross, J. C. (1987). The EAT and EDI: Screening, interrelationships, and psychometrics. *International Journal of Eating Disorders, 6*, 579-586.

Rivinus, T. M. (1987). Alcohol and drug abuse in college students. *Journal of College Student Psychotherapy, 1*, 5-31.

Rivinus, T. M., & Larimer, M. E. (1993). Violence, alcohol, other drugs, and the college student. *Journal of College Student Psychotherapy, 8*, 71-119.

Ross, H. E., & Tisdall, G. W. (1994). Alcohol use and abuse in a university psychiatric health serice: Prevalence and patterns of comorbidity with other psychiatric problems. *Journal of Alcohol and Drug Education, 39*, 63-74.

Wechsler, H., Dowdall, G. W., Davenport, A., & Castillo, S. (1995). Correlates of college student binge drinking. *American Journal of Public Health, 85*, 921-926.

Wechsler, H., Davenport, A., Dowdall, G., Moeykens, B., et al. (1994). Health and behavioral consequences of binge drinking in college: A national survey of students at 140 campuses. *Journal of the American Medical Association,* 272, 1672-1677.

Wernig, S. R. (1989). *Entering freshmen at Jesuit institutions, Fall 1988: A comparison of freshmen at nineteen Jesuit colleges and universities with all first-time, full-time freshmen.* Omaha, NE: Creighton University.

Author Note: J. M. Oliver, Department of Psychology, St. Louis University; Cynthia K. S. Reed, (now of Tarrant County College, Hurst, Texas); Bruce W. Smith, (now of Department of Psychology, Arizona State University).

Psychology and Religion: Indicators of Integration

Amber Haque
International Islamic University Malaysia

Although psychology and religion have experienced an uneasy relationship in the past, there is now a reversal of that trend seen from an increasing positive relationship between the two disciplines. In this paper, reasons for resurgence of psychology--religion dialogue are discussed in brief and a variety of indicators showing integration are examined. A list of relevant WebPages and e-mail discussion groups is also presented. Given the findings, it is concluded that the psychology--religion relationship will become stronger in the future and is perhaps gearing towards becoming an independent social scientific discipline in itself.

Despite attempts to integrate psychology and religion between the 19th Century and the mid 20th Century (Shafranske, 1996; Wulff, 1997), the two have been generally seen as mutually exclusive and the relationship between them viewed as one of paradox and impossibility. Based on an analysis of 50 psychological textbooks in the 1940s, Allport (1948) reported that in most psychological texts, "there is nothing to report except that they contain no treatment of the religious sentiment or closely related mental functions" (p.83). Several researches have also shown psychologist's antipathy toward religion to the extent that it became a taboo topic within psychology (Campbell, 1971; Kirkpatrick and Stone, 1935; Leuba, 1934). However, there now seems to be a reversal of the situation (Grace and Poelstra, 1995; Hefner, 1997; Miller and Martin, 1988; Shafranske, 1996; Watts and Williams, 1988) and a concern to rectify this situation is on the rise (Caplan, 1995; Haque, 1999; Loewenthal, 2000; Neelman and Persaud, 1995). There is even public interest in integrating the fields of psychology and religion (Peck, 1978; 1983). What is contributing to the resurgence of psychology--religion dialogue and what are the objective indicators of their integration? This paper briefly examines the possible reasons for resurgence and reviews the indicators of integration evident through an increase in professional organizations, university and institutional programs, academic journals, books and magazines, and even a fast developing Internet resource.

Author info: Correspondence should be sent to: Dr. Amber Haque, Dept. of Psychology, International Islamic U., Jalan Gombak, 53100 Kuala Lumpur, Malaysia.

North American Journal of Psychology, 2001, Vol. 3, No. 1, 61-76.

© NAJP

REASONS FOR RESURGENCE

There are situational as well as theoretical reasons contributing to the upsurge of the psychology--religion dialogue. Survey research indicates over 90% of the population in North America identifies itself with a religion and that religious beliefs have continuously grown since the American Revolution (Gallup and Jones, 1989; Shafranske, 1996). An increase of religious interest among the people would obviously lead psychologists to study religious phenomena whether or not psychologists profess to be religious themselves. Second, there is a great amount of research underway in the area of religion and mental health. Many psychologists now argue that religion is a potential positive factor in stabilizing the person's life (Bhugra, 1996; Comstock and Partridge, 1972; Gartner, 1996; Haque, 2000; Levin and Vanderpool, 1987; Zain and Varma, 1996). Third, the media also has portrayed religion as a controversial factor in the present day world, leading human science researchers to examine the influence of religion on human behavior (see for example, Pipes, 1990; Shaheen, 1999). Conflicts in the Middle East, Europe, and some other Asian countries in the name of religion attest to this fact. Fourth, perhaps the growth of a diversified psychology amidst a weakening influence of behaviorism on the American psychological scene has also given rise to alternative views of looking at human behavior. Fifth, a tremendous increase in the multicultural population in the North America with varying sets of belief systems has obligated helping professionals to consider a client's cultural and religious values in the evaluation and treatment process. This factor is tied to the growth of the scientist-practitioner model emphasizing practice-oriented approach where religion cannot be ignored as a personality variable. There is a related concern over the increased number of professionals who deal with "human" issues using a strictly "scientific" approach (Newman, 1989). Sixth, an impressive growth in the scientific inquiry of religion as a variable in human behavior has triggered further research in the field. Seventh, a more sophisticated yet less distinct psychology of religion in Europe and Asian countries where the tilt is toward phenomenological and transpersonal psychologies presents a challenge for similar development in North America (Wulff, 1997).

Within the theoretical reasons, first, there are psychologists who have challenged the limited domain of psychological research and argue that the metaphysical concepts, including soul, consciousness, transpersonal experience, and spirituality be studied in depth (Al Attas, 1990; Allport, 1950; Corbett, 1996; Fenn and Capps, 1995; Frankl, 1977; Helminiak, 1996; Jordan, 1995; Jung, 1933, 1938; Norager, 1996). Second, besides the views of the personality theorists, certain theoretical propositions of more recent origin regarding the influence of religion on humans are also responsible for this trend. For examples, consider the attributional theory of Spilka, Shaver and Kirkpatrick (1985); attachment theory of Kirkpatrick (1995) and the right-wing authoritarianism theory of Altemeyer and Hunsberger (1992). Third, a scientific undertaking of psychology and religion as a discipline needs a conceptual framework. The emphasis on measurement alone cannot fill the void of an adequate theory (Hood, Spilka, Hunsberger, and Gorsuch, 1996, p.23). Fourth, the post-modernism movement has also served as an impetus in reassessing psychology's evaluation of religion. Post-modernistic views emphasize that theories are merely constructs and theoretical truths are relative to personal and cultural contexts. Philosopher Paul Roth (1987) advises social scientists to consider a pluralistic view of human inquiry in which multiple methods can exist. Paloutzian (1996) elaborates this point further in explaining recent developments in the philosophy of science and psychological research methodology.

INDICATORS OF INTEGRATION

Professional Organizations

The earliest professional organizations that were established to emphasize integrated research include The American Society for Psychical Research (established 1885) and the Association for Clinical Pastoral Education (1925). Psychology and religion research in the mid 20th Century has resulted in the formation of the following organizations:

American Catholic Psychological Association (1947)
American Foundation of Religion and Psychiatry (1958)
American Nouthetic Psychological Association
Association for the Advancement of Psychosynthesis (1995)--dedicated to a psychological movement toward self-realization.
Association for Holotropic Breathwork International
Association for Religion and Intellectual Life
Association for the Scientific Study of Consciousness
Association for Transpersonal Psychology (1972)
Center for the Study of Psychiatry and Psychology (1975)
Christian Association for Psychological Studies (1953)
Forum on Religion, Spirituality and Aging (1990)
Hartford Institute for Religion Research
Association for Humanistic Psychology (1962)
International Transcendental Association
International Enneagram Association (1998)--combines modern psychology with ancient teachings as a psycho-spiritual tool for personal development)
International Federation for the Psychology of Religion, with its

secretariat in Belgium
International Muslim Psychological Association (1997)--based in Malaysia
Psychological Studies Institute (1971)
Religious Research Association
The National Academy of Religion and Mental Health (1954)
Society for the Scientific Study of Religion (1949)
Sufi Psychological Association (1996--transcends cultural boundaries and emphasizes the importance of soul and spirit

In 1976, a professional group called Psychologists Interested in Religious Issues (PIRI) achieved a division status with the American Psychological Association (APA) and later changed its name to the "Psychology of Religion," in 1993. According to this division, the psychology of religion "promotes the application of psychological research methods and interpretive frameworks to diverse forms of religion and spirituality; encourages the incorporation of the results of such work into clinical and other applied settings; and fosters constructive dialogue and interchange between psychological study and practice on the one hand and between religious perspectives and institutions on the other." This APA Division (36) publishes its own newsletter. APA even added "spirituality" to the index terms listed in the PsycLit computer database and currently has over 1000 items of readings on spirituality itself.

The APA now encourages its members to view religion as a significant aspect of human life. The 1992 APA Code of Conduct also specifies that in the absence of such a service from the psychologist, appropriate referrals for the clients should be made to ensure proper and complete treatment of the patients. Considering the importance of religion, the Diagnostic and Statistical Manual published by the American Psychiatric Association, included in its 4th edition (1994) a classification on "religious problems" which further indicates a growing recognition among the scientific community of religion as a factor in mental health.

Integrated Programs

In the West, the first integrated and APA accredited doctoral program in psychology was offered at the Fuller Theological Seminary in 1988. Biola University (Rosemead Graduate School of Psychology) opened its doors for students in 1970. Emory University offers a Ph.D. in Religion and Personality, and Columbia University offers a joint master's degree in Religion and Social work. Doctoral programs are also offered at Baylor University (Texas), George Fox University (Oregon), and Wheaton College (Illinois). The University of Tennessee at Martin also offers an integrated program in the Department of Psychology, Philosophy, and Religious Studies. There are even online psychology and religion courses offered as part of the Bachelor's of Theology Program at Canyon College (Idaho). The University of Wales at Bangor (UK) offers advanced psychology of religion courses at its Center for Ministerial Studies. Integrated graduate programs are also offered at the John F. Kennedy Graduate School for Holistic Studies, Greenwich University (Hawaii), Naropa University (Colorado), and Trinity Western University (Canada).

Apart from the traditional religious programs, transpersonal psychology is offered in more than 130 institutions that include fully accredited and non-accredited programs (Vande Kemp, 1996, p. 81). The Psychology Department of the State University of West Georgia publishes a Directory of Graduate Programs in Humanistic-Transpersonal Psychology in North America and is available online at http://www.sonoma.edu/projects/os2/grad/os2gradschool.htm. Other religion related integration includes the Consciousness Research and Training Project, Psychosynthesis Project, and Spiritual Direction and Formation, all listed in the Common Boundary Graduate Education Guide (Demetrios, Simpkinson, and Bennet, 1991). The Haden Institute in Charlotte (North Carolina) offers courses dealing with psychology and religion and the ISIS Institute in El Sobrante (California) focuses on Transpersonal Psychology. The Woodfish Institute based in San Francisco (California), advocates integration of ancient healing methods and modern psychology based on the practices of Eco-psychology and Shamanism (the oldest form of mind-body healing method practiced through altered states of consciousness). In the 1970's, a new research area called "Quality of Life," which encompasses psychology and religion, also began to emerge (Ellison and Smith, 1991; Moberg and Brusek, 1978; Paloutzian and Ellison, 1983).

Paloutzian (1996) suggests that a more discreet way to judge whether religion is getting the deserved status within psychology is to see if religion is taught within the Introduction to Psychology course, as the authors would include most important topics in their field at the time of their writing. In their research, Lehr and Spilka (1989) found that there was substantial progress in the 1980's compared to the 1970's in using religious contents in introductory textbooks of psychology. An earlier study by Spilka, Comp, and Goldsmith (1981) also compared references to religion in introductory textbooks of psychology between 1950s and 1970s. Although slightly less percentage of citations in the 1970s compared to the 1950s were found, the report indicated that beginning in the 1970s, religion was covered in a more "objective" manner and empirical research started to increase.

In a comparison of the editions of *Psychology and Life*, the oldest textbook in general psychology, Paloutzian (1996, p.51) reported an upsurge of religious content in the 9th and 10th editions (1937-1979), but the sections on religious content were missing in the 11th through the 13th editions. Loewenthal (2000) pointed out that in the first half of the 1990s there was a 400% increase in mean references to religion per book in undergraduate social psychology and psychology of personality textbooks (p. 10). Also the annual number of publications in religion and mental health went up by 400%-500% in the 1990s compared to the 1980s (Dein and Loewenthal, 1998).

A review of the existing literature suggests that Christian graduate training programs and professional societies in psychology have developed more rapidly compared to integrated training programs for other religions of the world. The reason for more integrated programs in Christian countries may be attributed to the availability of funds, hard work, and organized efforts, as well as greater marketability of those programs. Another reason for the growth of such programs could be the absorption of graduates into Christian counseling centers, hospitals and other religiously oriented human service agencies. Yet another reason for apparent rapid growth of integrated programs in the West and largely Christian world is a lack of research available from other communities and the Eastern world.

Several places that offer integrated postgraduate programs in the East are Brokenshire College in Davo City, Philippines, where programs in the area of psychology and religion are based on the Christian faith; International Islamic University of Malaysia, where integrated programs were established in the mid 1990s in its various human sciences disciplines including psychology, and The International Institute of Islamic Thought and Civilization, in Kuala Lumpur, Malaysia, developed similar programs in the early 1990s. Other programs available in Asia are mostly in local languages and are not publicized well. The integration of religion and psychology has also taken place in Jewish (Meier, 1988), other Muslim (Rizvi, 1988), and Buddhist (De Silva, 1996) communities, but their numbers are very few (Tan, 1996).

Academic Journals

Two journals that integrated psychology and religion were established at the beginning of the 20th century, i.e. *The American Journal of Psychology and Education* (1904) and the *Journal of American Society for Psychical Research* (1907). A majority of the integrated psychology journals emerged in the latter half of the twentieth century, especially in the 1980s and 1990s, for example, *The Journal of Psychology and*

Theology established in 1973, *The Journal of Psychology and Christianity* established in 1982, *The Journal of Psychology and Judaism*, and *The International Journal for the Psychology of Religion*, established in 1990. *The International Journal for the Psychology of Religion* is a platform for international psychologists interested in this area and presents a wide array of articles, essays and reviews from Australia (O'Connor, 1996; 1998), Brazil (Carvalho and Sato, 1997), Canada (Hunsberger, Alisat, Pancer, and Pratt, 1996; Ross, Weiss, and Jackson, 1996), Denmark (Norager, 1996), England (Argyle and Hill, 2000; Kose and Loewenthal, 2000; Watts, 1996), France (Dechocny, Hurtean, Quelen, and Ragot, 1997), Germany (Hermsen, 1996), Iceland (Haroldsson and Houtkooper, 1996), Italy (Aletti, 1992), The Netherlands (Belzen, 1994; 1996), Poland (Grzymala-Moszczynska, 1991), Scandinavia (Wikstrom, 1993), Sweden (Geels, 1996; Granqvist and Hagekull, 2000), and Turkey (Kose, 1996).

Other integrated journals include, *Journal of Religion and Health* established in 1961, *Journal for the Scientific Study of Religion* (1961), *The Journal of Transpersonal Psychology* (1968), *Journal of Humanistic Psychology, Review of Religious Research* (1969), *Journal of Religious Gerontology* (1989), *Research in the Social Scientific Study of Religion, Pastoral Psychology, The International Journal of Transpersonal Studies* (1993--previously known as *The Australian Journal of Transpersonal Psychology* established in 1981), *The Mind's Eye Journal* (1996), and the *Sufi Psychology Association Journal* (1996).

The Association for the Scientific Study of Consciousness publishes two official journals, i.e. 1) *Consciousness & Cognition* (1994) and 2) *Psyche* (1994). *Cross Current* is another psychological journal focusing on integration of the soul and the mind. The Association of Muslim Social Scientists in the USA started an integrated journal called the *American Journal of Islamic Social Sciences* (AJISS), in 1973. AJISS published a special "Psychology Issue" in 1998, where several areas of general psychology were addressed from an Islamic religious perspective. *The Intellectual Discourse* (1992) published from the International Islamic University of Malaysia is another journal of integrated nature as well as the *Journal of Islamic Studies* published from UK and *Hamdard Islamicus* (1982) published from Pakistan.

Books and Magazines

A few books that appeared at the start of the 20th Century are Edward Ames' *Psychology of Religious Experience* (1910), George Stratton's *Psychology of Religious Life* (1911) followed by James Pratt's *The Religious Consciousness* (1920). A few more books appeared in Europe, e.g. *An Introduction to the Psychology of Religion* by Robert Thouless (1923) and the *Psychology of Religious Mysticism*, by James Leuba (1925). Then again, numerous books started to appear on the subject in the 1980s and 1990s. Authors of these works include Batson and Ventis, 1982; Belzen, 2000; Brown 1987, 1994; Capps, Rambo, and Ransohoff, (1976); Corbett, 1996; Crapps, 1986; Fenn and Capps, 1995; Hood, Spilka, Hunsberger, and Gorsuch, 1996; Jeeves, 1997; Loewenthal, 2000; Malony, 1991; Meadow and Kahoe, 1984; Paloutzian, 1996; Pargament, 1997; Shafranske, 1996; Spilka, Hood, and Gorsuch, 1985; Spilka and McIntosh, 1996; Thouless, 1971; Vande Kemp, 1984; and Wulff, 1997. A special edition of *The Journal of Psychology and Theology* (Grace and Poelstra, 1995) deserves mention here as well, as it relates to exploring the integration of psychology and religion in undergraduate psychology courses. Most of these books are written by Christian psychologists and published in the West.

Books written on psychology and religion from Asian perspectives are also available, and the number is growing (for details, see Kao & Sinha, 1997; Kuppuswamy, 1993; Singh, 1979; Wallace, Orme-Johnson, and Dillbeck, 1993). Some references integrating Islam and Psychology are also available in a book by Yasien Mohamad (1998) and in an earlier treatise on *A History of Muslim Philosophy* by M.M. Sharif (1963). Zafar A. Ansari (1992) also compiled papers on integrating psychology and religion presented in an Islamic conference held in Islamabad, Pakistan. Many books are available on the general subject of psychology and religion but in local languages; for example, a significant amount of research has been conducted in Arabic-speaking countries, but the English translation of such works is not currently available (Haque, 1999).

Monographs, chapters, and articles on psychology and religion have also appeared in many places, i.e., *The Annual Review of Psychology* (Gorsuch, 1988), *American Psychologist* (Bergin, 1991; Jones, 1994; Weiner, 1993) and *Psychological Science* (Sethi & Seligman, 1993; Waller, Kogetin, Bouchard, Lykken, and Tellegen, 1990).

There are magazines as well that relate to the psychology--religion interface, e.g. *Alphaomega*, published from Sweden, addresses psychology and spirituality issues. *Highest Vision*, published from La Jolla (California) is a metaphysical magazine emphasizing understanding and applications of spiritual principles in everyday life; *Omniscience* published in Venice (California) explains spiritual psychology in order to understand the self; *Spirituality and Health* published from New York, contains articles about soul-body connection; and *Science and Spirit* published from Oxford (UK), emphasizes integration of spiritual and scientific theories. Certain publishers also specialize in religious books from a psychological perspective, e.g. Larsen publications (New York) and Oneworld Publications (distributed by Penguin Books Worldwide). Binc Books (Oklahoma City) specializes in rare books in the psychology and religion area.

Psychology of Religion Web Pages

The following web pages are created by individuals or organizations interested in psychology and religion issues. Certain pages also provide links to other similar Internet sites.

Http://www.psychwww.com/psyrelig/index.htm: [maintained by Michael E. Nielsen, Ph.D., Georgia Southern University, USA]

Http://www.socsci.kun.nl/psy/cultuur/users/jablonski/psycherel.html: [established by Przemyslaw Jablonski, University of Nijmegen, Netherlands]

Http://www.mdani.demon.co.uk/: [maintained by Michael Daniels of the Center for Applied Psychology, Liverpool John Moore University]

Http://www.vanguard.edu/psychology/webintegration.html: [maintained by Douglas Degelman, Ph.D. Professor of Psychology at Vanguard University of Southern California]

Http://www.comp.glam.ac.uk: [Website of the Catholic Psychology Group, who meet twice a year to discuss a wide range of topics of common interest]

Http://www.ozemail.com.au/~cohnmat/bpsy.html: [created by Matthew Cohn - includes, biblical psychology e-text library, questions and issues on biblical psychology]

Http://www.psy.Kuleuven.ac.be/religion/index/.html [Catholic University, Belgium]

Http://www.home.cc.umanitoba.ca/~creamer/homepage/index.html [maintained by Dr. Creamer, Professor of Psychology, University of Manitoba, Canada]

Http://www.angelfire.com/al/islamic_psychology/ [views on psychology from Islamic perspective]

http://www.ummah.net/bicnews/Articles/forgotten.htm [emphasizes integration of psychology and religion from Islamic point of view]

http://www.aristotle.net/~crazybob/ [Philosophy--psychology--religion debate, psychology online, personality tests]

Http://sufi-psychology.org/info_req.html [emphasizes the importance of soul and spirit--based in California]

E-mail Discussion Forums

The following discussion forums encourage scholarly discourse on psychology and religion issues through electronic input. Each site has

BEST COPY AVAILABLE

specific research interests and some have discussion forums for its members only.

Http://www.lightlink.com/xine/psy_religion/psyreltoc.html [psyrel-l is a list for the academic discussion of psychology of religion, its history, focus, cultural milieu, methods, and approaches. Discussion is encouraged across a broad spectrum but primarily of academic nature. Maintained by Christine Norstrand (xine@lightlink.com) and Greg Singleton]

Http://www.coolboard.com/msgshow.cfm/msgboard [general psychology of religion discussion zone]

Http://www.angelfire.com/az/psycrelig/ [general psychology of religion discussion zone]

Http://www.egroups.com.hk/dir/Science/Social_Sciences/Psychology/Psychology_of_Religion [pcr-list is intended to serve members of the person, culture & religion group of the American Academy of Religion and other interested persons as a forum for discussion of issues related to shared concern with the relationships between religion, psychology, and contemporary cultures]

Http://www.egroups.com/group/transpersonal_psychology [discussion of transpersonal psychology as taught by authors Michael Washburn, Stanislav Grof, John E. Nelson, Ken Wilber, Charles Tart, etc.]

Http://www.aap-psychosynthesis.org/cullen.html [discussion focuses on theory, therapy, case studies, managerial psychosynthesis, group psychosynthesis, organizational psychosynthesis, typology, research, intuition, creativity, education, etc.]

CONCLUSION

A literature review of the relationship between psychology and religion suggests that very few integrated professional organizations and journals existed before the middle of the 20th Century. Substantial growth started to take place in the 1970s and significant increase started in the 1980s and 1990s. This change, although more apparent in North America, is also noted in other countries through indicators such as burgeoning professional organizations and institutes, university programs, academic journals, books and magazines, and the growing number of Internet sites. This review is by no means inclusive, and further research in collaboration with psychologists from other parts of the world would be highly beneficial in measuring the international growth and trends of the psychology--religion interface. The present review suggests that the integration of psychology and religion as a discipline is well on its way toward becoming an independent area of social science research.

REFERENCES

Al Attas, S.M.N. (1990). *The nature of man and the psychology of the human soul.* Kuala Lumpur: International Institute of Islamic Thought and Civilization.

Aletti, M. (1992). The psychology of religion in Italy. *The International Journal for the Psychology of Religion, 2,* 171-189.

Allport, G.W. (1948). *Psychology*. In College Reading and Religion. New Haven, Conn.: Yale Univ. Press.

Allport, G. W. (1950). *The individual and his religion*. New York: Macmillan.

Altemeyer, B. and Hunsberger, B. (1992). Authoritarianism, religious fundamentalism, quest, and prejudice. *International Journal for the Psychology of Religion, 2,* 113-133.

American Psychological Association (1992). Ethical principles of psychologists and code of conduct. *American psychologist, 47,* 1597-1611.

Ames, E.S. (1910). *The psychology of religious experience*. Boston: Houghton Mifflin.

Ansari, Z.A. (Ed.) (1992). *Quranic concepts of human psyche*. Lahore, Pakistan: International Institute of Islamic Thought.

Argyle, M. and Beit-Hallahmi, B. (1975). *The social psychology of religion,* London: Routledge.

Argyle, M. and Hill, P. (2000). Religious experiences and their relations with happiness and personality, *The International Journal for the Psychology of Religion, 10, 3,* 157-172.

Batson, C.D. and Ventice, W.L. (1982). *The religious experience: a social-psychological perspective*. New York: Oxford University Press.

Belzen, J.A. (1994). Between fest and famine: A sketch of the development of the psychology of religion in The Netherlands. *The International Journal for the Psychology of Religion, 4,* 181-197.

Belzen, J.A. (1996). Beyond a classic? Hjalmer Sunder's Role theory and contemporary narrative psychology, *The International Journal for the Psychology of Religion, 6, 3,* 181-199.

Belzen, J. A. (Ed.) (2000). *Aspects in contexts: Studies in the history of psychology of religion.* Amsterdam/Atlanta: Rodopi.

Bergin, A.E. (1991). Values and religious issues in psychotherapy and mental health, *American Psychologist, 46,* 394-403.

Bhugra, D. (1996). *Psychiatry and religion: Context, consensus and controversies*, London: Routledge.

Brown, L.B. (1987). *The psychology of religious belief.* London: Academic Press.

Brown, L.B. (1994). *The human side of prayer: The psychology of praying.* Birmingham, AL: Religious Education Press.

Campbell, D.P. (1971). *Handbook for the strong vocational interest blank,* Calif.: Stanford Univ. Press.

Caplan, P. (1995). *They say you're crazy.* Addison-Wesley Publishing.

Capps, D., Rambo, L., and Ransohoff, P. (1976). *Psychology of Religion: A guide to information sources*. Detroit: Gale Research Co.

Carvalho, A.M. and Sato, T. (1997). Comparing religious identity of Baptist and Presbyterian Bible college students by minimum spanning tree. *The International Journal for the Psychology of Religion, 7,2,* 111-125.

Comstock, G.W. and Partridge, K.B. (1972). Church Attendance and Health. *Journal of Chronic Disease, 25,* 665-672.

Corbett, L. (1996). *The religious function of the psyche*. London: Routledge.

Crapps, R.W. (1986). *An introduction to the psychology of religion*. Macon, GA: Mercer University Press.

Dechocny, J-P, Hurtean, C., Quelen, F. and Ragot, I. (1997). The psychology of religion and cognitive models, *The International Journal for the Psychology of Religion, 7, 4,* 263-268.

Dein, S. and Loewenthal, K.M. (1998). *Editorial. Mental Health, Religion, and Culture, 1,* 5-9.

Demetrios, E. Simpkinson, C.H. and Bennet C. (Eds.). (1991). *The Common Boundary Graduate Education Guide.* Bethesda, MD: Common Boundary.

DeSilva, P. (1996). *Buddhist psychology and implications for treatment. In D. Bhugra (Ed.) Psychiatry and Religion: Context, Consensus, and Controversies*. London: Routledge.

Diagnostic and Statistical Manual of Mental Disorders (1994). Washington D.C.: American Psychiatric Association.

Ellison, C.W. and Smith, J. (1991). Toward an integrative measure of health and well being. *Journal of Psychology and Theology, 19,* 35-48.

Fenn, R. K. and Capps, D. (Eds.) (1995). *On losing the soul: Essays in the social psychology of religion*. State University of New York Press.

Frankl, V. (1977). *The unconscious God*. London: Hodder & Stroughton.

Gallup, G. and Jones, S. (1989). *One hundred questions and answers: Religion in America*. Princeton, NJ: Princeton Religion Research Center.

Gartner, J. (1996). *Religion and mental health. In E.P Shafranske (Ed.) Religion and the clinical practice of psychology*, pp. 187-214. Washington D.C.: American Psychological Association

Geels, A. (1996). A Note on the psychology of Dhikr: The Halveti-Jerrahi order of dervishes in Istanbul. *International Journal for the Psychology of Religion, 6, 4,* 229-251.

Gorsuch, L. R. (1988). Psychology of religion, *Annual Review of Psychology, 39,* 201-221.

Grace, C.R. and Poelstra, P.L. (Eds.), (1995). *Journal of Psychology and Theology, Special Issue: An Exploration of integrative issues in undergraduate psychology programs.*

Granqvist, P. and Hagekull, B. (2000). Religiosity, adult attachment and why singles are more religious, *International Journal for the Psychology of Religion, 10, 2,* 111-123.

Grzymala-Moszczynska, H. (1991). The psychology of religion in Poland. *International Journal for the Psychology of Religion, 1,* 243-247.

Haque, A. (1999). Psychology and Religion: Their relationship and integration from Islamic perspective. *American Journal of Islamic Social Sciences, 15, 4,* 97-116.

Haque, A. (2000). Psychology and religion: Two approaches to positive mental health. *Intellectual Discourse, 8,1,* 81-94.

Haroldsson, E.H. and Houtkooper, J.M. (1996). Traditional Christian beliefs, spiritualism, ands the paranormal: An Icelandic-American comparison. International *Journal for the Psychology of Religion, 6,1,* 551-64.

Hefner, P. (1997). The science-religion relation: controversy, convergence, and search for meaning. *International Journal for the Psychology of Religion, 7,* 143-158.

Helminiak, D.A. (1996). A scientific spirituality: The interface of psychology and religion. *International Journal for the Psychology of Religion, 6,1,*1-19.

Hermsen, H.E. (1996). Person-centered psychology and Taoism: The perception of Lao-tzu by Carl R. Rogers. *The International Journal for the Psychology of Religion, 6,2,* 107-125.

Hood, R. W., Spilka, B., Hunsberger, B. and Gorsuch, R. (1996). *The Psychology of Religion: An Empirical Approach (2nd Ed).* New York: The Guilford Press.

Hunsberger, B. Alisat, S. Pancer, S.M., Pratt, M. (1996). Religious fundamentalism and religious doubts: content, connections, and complexity of thinking, *International Journal for the Psychology of Religion, 6, 3,* 201-220.

Jeeves, M. A. (1997). *Human Nature at the Millennium: Reflections on the integration of psychology and Christianity.* Baker Book House.

Jones, S.L. (1994). A constructive relationship for religion with the science and profession and psychology: Perhaps the boldest model yet. *American Psychologist, 49, 3,* 184-199.

Jordan, N. (1995). *Themes in speculative psychology. In David Cohen (Ed.), Psychologists on Psychology.* New York: Routledge.

Jung, C.G. (1933). *Modern man in search of a soul.* New York: Harcourt Brace.

Jung, C.G. (1938). *Psychology and religion. In Collected Works, 11, 2,* 1969, pp. 3-105.

Kao, S.R. and Sinha, D. (1997). *Asian Perspectives on Psychology,* New Delhi: Sage Publications.

Kirkpatrick, L.A. (1995). *Attachment theory and religious experience. In R.W. Hood, Jr., Handbook of Religious Experience,* pp. 446-475. Birmingham, AL: Religious Education Press.

Kirkpatrick, C. and Stone, S. (1935). Attitude measurement and the comparison of generations. *Journal of Applied Psychology, 5,* 564-582.

Kose, A (1996). Religious conversion: is it an adolescent phenomena? The case of native British converts to Islam. *International Journal for the Psychology of Religion, 6, 4,* 253-262.

Kose, A, and Loewenthal, K.M. (2000). Conversion motifs among British converts to Islam, *International Journal for the Psychology of Religion, 10,* 2, 101-110.

Kuppuswamy, B. (1993). *Source Book of Ancient Indian Psychology*, Delhi: Konark Publishers.

Lehr, E. and Spilka, B. (1989). Religion in the introductory psychology textbook: a comparison of three decades, *Journal for the Scientific Study of Religion, 28*, 366-371.

Leuba, J.H. (1925). *The psychology of religious mysticism*. New York: Harcourt Brace.

Leuba, J.H. (1934). Religious beliefs of American Scientists. *Harper's Magazine, 169*, 291-300.

Levin, J. and Vanderpool, H. (1987). Is Frequent Religious Attendance Really Conducive to Better Health?: Toward an Epidemiology of Religion. *Social Science Medicine, 24*, 589-600.

Loewenthal, K.M. (2000). *The psychology of religion: A short introduction*. Oxford: Oneworld.

Malony, H. N. (Ed.). (1991). *Psychology of religion: Personalities, problems, possibilities*. Grand Rapids, MI: Baker Book House.

Meadow, M.J. and Kahoe, R.D. (1984). *Psychology of Religion: Religion in Individual Lives*. New York: Harper & Row.

Meier, L (1988). *Jewish values in psychotherapy*. New York: University Press of America

Miller, W.R. and Martin, J.E. (1988). *Behavior therapy and religion: Integrating behavioral approaches to change*. Newbury Park: Sage Publications.

Moberg, D.O. and Brusek, P.M. (1978). Spiritual well being: A neglected subject in quality of life research, *Social Indicators Research, 5*, 303-323.

Mohamed, Y. (1998). *Human nature in Islam*. Kuala Lumpur: A.S. Noordeen.

Neelman, J and Persaud, R. (1995). Why do psychiatrists neglect religion? *British Journal of Medical Psychology, 68*, 169-178.

Newman, J.K. (1989). A theological perspective on the licensing of helping professionals. *Journal of Psychology and Theology, 17*, 252-262.

Norager, T (1996). Metapsychology and dissonance: a note on some neglected issues on psychology of religion. *International Journal for the Psychology of Religion, 6, 3*, 139-149.

O'Connor, K.V. (1996). Reconfiguring the emotions in the psychology of religion, *International Journal for the Psychology of Religion, 6, 3*, 165-173.

O'Connor, K.V. (1998). Religion and mental health: A review of Antoine Vetgote's approach in guilt and desire, *International Journal for the Psychology and Religion, 8*, 125-148.

Paloutzian, R.F. and Ellison, C.W. (1983). *Loneliness, spiritual well-being and quality of life. In L.A. Peplan and D. Perlamn (Eds.), Loneliness: A Sourcebook of Current Theory Research and Therapy*. New York: Wiley Interscience.

Paloutzian R.F. (1996). *Invitation to the Psychology of Religion (2nd ed.)*. Boston: Allyn & Bacon.

Pargament, K. I. (1997). *The psychology of religion and coping: Theory, research, practice*. New York: Guilford.

Peck, M.S. (1978). *The road less traveled*. New York: Simon and Schuster.

Peck, M.S. (1983). *People of the lie: The hope for healing human evil.* New York: Simon and Schuster.

Pipes, D. (1990) The Muslims are coming, The Muslims are coming. *National Review, Nov.19,* 28-31.

Pratt, J.B. (1920). *The religious consciousness: A psychological study.* New York: Macmillan.

Rizvi, S.A.A. (1988). *A Muslim tradition in psychotherapy and modern trends.* Lahore, Pakistan: Institute of Islamic Culture.

Ross, C.F.J., Weiss, D. and Jackson, L (1996). The relation of Jungian psychological type to religious attitudes and practices, *International Journal for the Psychology of Religion, 6 (4),* 263-269.

Roth, P.A. (1987). *Meaning and method in the social sciences: A case for methodological pluralism. Ithaca: Cornell University Press. In D.M. Wulff (1997), Psychology of Religion: Classic and contemporary.* Mass: John Wiley & Sons.

Sethi, S. and Seligman, M.E.P. (1993). Optimism and fundamentalism, *Psychological Science, 4,* 256-259.

Shafranske, E.P. (Ed.) (1996). *Religion and the clinical practice of psychology.* Washington D.C.: American Psychological Association.

Shaheen, J. (1999). *Hollywood's Reel Arabs and Muslims. In A. Haque (ed.), Muslims and Islamization in North America: Problems and Prospects,* pp. 179-202, MD: Amana Publications.

Sharif, M.M . (1963). *A history of Muslim philosophy.* Karachi, Pakistan: Royal Books.

Singh, A.K. (1979). *Development of religious identity and prejudice in Indian children. In A. De Souza (ed.), Children in India: Critical Issues in Human Development.* New Delhi: Manohar Press.

Spilka, B., Comp, G., and Goldsmith W.M. (1981). Faith and behavior: Religion in introductory texts of the 1950s and 1970s', *Teaching of Psychology, 8,* 159-160.

Spilka, B., Shaver, P., and Kirkpatrick, L.A. (1985). A general attribution theory for the psychology of religion. *Journal for the Scientific Study of Religion, 24,*1-20.

Spilka, B., Hood, R.W., Jr., and Gorsuch, R.L. (1985). *The psychology of religion: An empirical approach.* Englewood Cliffs, NJ: Prentice Hall.

Spilka, B. and McIntosh, D. N. (Eds.) (1996). *The psychology of religion: Theoretical approaches.* Bounder, CO: Westview.

Stratton, G.M. (1911). *Psychology of the religious life.* London: George Allen and Unwin.

Tan, S.Y. (1996). *Training in professional psychology: diversity includes religion. In E. P. Shafranske (Ed.), Religion and clinical practice of psychology.* Washington D.C.: American Psychological Association.

Thouless, R.H. (1923). *An introduction to the psychology of religion.* London: Cambridge University Press.

Thouless, R.H. (1971). *An introduction to the psychology of religion (3^{rd} Ed.).* London: Cambridge University Press.

Vande Kemp, H. (1984). *Psychology and theology in western thought 1672-1965: A historical and annotated bibliography*. New York: Kraus International Publications.

Vande Kemp, H. (1996). *Historical perspective: religion and clinical psychology in America. In E. P. Shafranske (Ed.), Religion and clinical practice of psychology*. Washington D.C.: American Psychological Association.

Wallace, R.K., Orme-Johnson, D.W. and Dillbeck, M.C. (1993). *Scientific Research on Maharishi's Transcendental Meditation Program: Collected Papers* (Vol.5), Fairfield, IA: Maharishi International University Press.

Waller, N.G., Kogetin, B.A., Bouchard, T.J., Lykken, D.T., and Tellegen, A. (1990). Genetic and environmental influences on religious interests, attitudes, and values: A study of twins reared apart and together. *Psychological Science, 1*, 2, 138-142.

Watts, F. and Williams, M. (1988). *The psychology of religious knowing*, New York: Cambridge University Press.

Watts, F (1996). Emotion and Religion Revisited: A response to O'Connor and Averill, *International Journal for the Psychology of Religion, 3*, 175-179.

Weiner, B. (1993). On sin versus sickness: A theory of perceived responsibility and social motivation. *American Psychologist, 48, 9*, 957-965.

Wikstrom, O. (1993). The psychology of religion in Scandinavia. *International Journal for the Psychology of Religion, 3*, 47-65.

Wulff, D. M. (1997). *Psychology of religion: Classic and contemporary (2nd ed.)*. New York: John Wiley.

Zain, A.M. and Varma, S.L. (1996). Religious Psychotherapy--A Proposed Model Based on The Malaysian Experience, *Journal of FIMA, 1*, 118-123.

Descriptions of Religious Experience among U.S. Christians and Malaysian Muslims

Michael E. Nielsen
Georgia Southern University

Rick Stevens
University of Louisiana at Monroe

Cross-cultural descriptions of religious experiences were investigated in two studies using adjectives derived from the five-factor model of personality. In study 1, 93 U.S. Christians indicated that adjectives reflecting low levels of neuroticism, and high levels of agreeableness and conscientiousness reflected how they felt during religious experiences. Adjectives reflecting positive affect also were rated to be descriptive of religious experiences. Study 1 also found that participants do not have religious experiences as frequently as they would like. In study 2, 62 Malaysian Muslims were asked to describe either their most important religious or spiritual experience. These descriptions followed a similar pattern as was found in Study 1, with participants describing themselves as being more emotionally stable, open and extraverted during religious than during spiritual experiences. Results are discussed in terms of the utility of adjective ratings for describing experiences across diverse religious traditions.

A fascinating area in psychological studies of religion concerns people's descriptions of religious experiences. In studies of religious experience, as in studies of many psychological phenomena, researchers rely on self-reports because the phenomenon cannot be independently and objectively confirmed. Researchers therefore must determine the method most suitable for obtaining self-reports. Most often, the method has been to devise a scale of statements based on the writings of religious figures or mystics for describing the experiences (e.g., Hood, 1975; Spilka, Brown & Cassidy, 1992).

Such an approach has the advantage of strong face validity. The mystic or religious leader, who ostensibly is quite familiar with the phenomena, may be better suited than most people to articulate the characteristics of an experience characterized by its very ineffability. At the same time, the approach may possess certain disadvantages. The applicability of these descriptions to the average person's religious

---

Author info: Correspondence should be sent to: Michael Nielsen, Department of Psychology, Georgia Southern University, Statesboro, GA 30460-8041, email mnielsen@gasou.edu.

North American Journal of Psychology, 2001, Vol. 3, No. 1, 77-90.

© NAJP

experience may be questioned. Religious leaders and mystics differ from their followers and non-mystics in many ways, and the experiences that each group has are conceivably quite different in nature, intensity, and/or frequency. For example, although 84% of people in the U.S. state that religion is important (Gallup & Jones, 1989), research also indicates relatively low rates of mystical experience by people, and that simple surveys may actually overestimate the rate at which mystical experiences occur (Thomas & Cooper, 1978, 1980). An additional disadvantage is the extent to which such scales may be limited to a single religious tradition. As an examination of Hill and Hood's (1999) compendium of scales reveals, the majority of scales include content specific to one religion, usually Christianity. The culture-bound content of scales reduces their utility across religious or spiritual traditions.

Several alternatives exist to studying religious or mystical experience via surveys. One is to examine experiences in terms that are more broadly applied to the description of people or events in general. Such an approach will inform us of broad ways that religious experiences might be characterized, and perhaps distinguish religious experiences from other kinds of experiences. This is the approach taken in the present studies, which explore how general adjectives might be used to describe religious experiences. The logic underlying this lexical approach is based on Allport and Odbert's (1936) classic study, which assumed that the terms found in everyday language are assumed to be useful in describing people and their experiences. Research in this vein has focused on personality, and, to a lesser extent, affect.

Many lexical studies of personality have revealed five factors or dimensions: neuroticism (emotional stability), extraversion, openness, agreeableness, and conscientiousness (McCrae & Costa, 1985, 1987). McCrae and Costa's (1985, 1987) work on the five factor model is interesting in that it demonstrated substantial correlations between simple adjective pairs intended to measure the five factors (the lexical tradition) and more elaborate scales developed for that purpose. Their research, involving more than 1,000 participants, also demonstrated a high correlation between self- and other-ratings on the five factors. This suggests that people's perceptions of themselves are consistent with others' perceptions. To the extent that self-ratings may be valid, then, the use of adjectives represents one alternative to researchers seeking to understand how people view religious experiences.

Researchers have begun to investigate the connections between religious experience and the five-factor model of personality. MacDonald (2000) investigated several dimensions of spirituality, and their correlation with the five-factor model of personality. Spirituality was assessed in five dimensions: cognitive, experiential, existential well-being, paranormal beliefs, and religiousness. Analyses revealed that neuroticism correlated negatively with existential well-being. Extraversion correlated positively with all spirituality dimensions except paranormal beliefs. Openness correlated positively with cognitive, experiential, and paranormal beliefs. Agreeableness and conscientiousness correlated positively with cognitive, existential well-being, and religiousness. A second recent study using this approach asked participants to use adjectives representing the five-factor model of personality to describe how participants felt during religious experiences (Nielsen, 2000). Similar to the results found by MacDonald (2000), people described their feelings during religious experiences as being high in agreeableness and low in neuroticism.

The following two brief studies seek to extend this research in several ways. In the first study, we examine the utility of affect adjectives for describing how participants feel during a religious experience. That study also assesses the extent to which participants' religious experiences match ideal or model religious experiences. In Study 2, we consider the descriptions of religious experiences among Malaysian Muslims, and the extent to which spiritual experiences are distinct from religious ones.

STUDY 1

Notable for its absence in research regarding religion is Russell's (1980, 1983) two-dimensional taxonomy of affect. The first dimension, positive vs. negative affect[1], is characterized by adjectives such as pleased and miserable. The second dimension, intensity, is characterized by adjectives such as aroused and sleepy. Crossing these two dimensions yields four combinations of arousal and affect: high-intensity positive (joyful); high-intensity negative (afraid); low-intensity positive (at-ease); and low-intensity negative (bored).

There would appear to be many possibilities to apply this model of affect to religious experience. The present study begins by examining the perceived relevance of these dimensions to religiousness. Specifically, how well do adjectives corresponding to these dimensions reflect religious experience? Because most Americans consider religion to be "very important" (Gallup & Jones, 1989), and based on the assumption that people generally seek pleasure and avoid pain, it was hypothesized that adjectives containing a positive valence would be considered more descriptive of religious experience than would negatively-valenced adjectives. No prediction was made regarding the intensity dimension.[2]

Study 1, then, represents an attempt to replicate and extend the findings of Nielsen (2000). Participants rated words, derived from personality and affect taxonomies, in terms of how well each word described how they felt during their religious experiences. Among the personality adjectives, words representing low neuroticism (emotional stability) and high agreeableness were expected to receive the highest ratings. Among affect adjectives, words representing positive valence were expected to receive higher ratings than negatively valenced words.

The congruence between a word's importance in describing religious experiences, and how often the word actually is experienced, also was investigated. In a sense, this is similar to Rogers' notion of "Real" vs. "Ideal" self-concepts (Rogers, 1976). Participants were asked to indicate both how important an adjective is for describing religious experiences, and how often they actually experience that adjective in their religious lives. This was done in order to explore possible discrepancies between how people would like their religious experiences to be, and how the experiences actually are.

Method

Participants. The snowball method of sampling (Henry, 1990) was used for Study 1. Participants were 93 Illinois adults not enrolled in school. Mean age of participants was 39 (*S.D.* = 13.1, range 16 to 87).[3] Participants tended to be female (52%), reported attending church services an average of 2.9 times during the previous month (*S.D.* = 3.7, range 0 to 25), and reported praying an average of 7.9 times during the previous week (*S.D.* = 8.2, range 0 to 45). The most common denominational affiliations were Catholic (*N* = 26), Lutheran (9), Mormon (8), Presbyterian (6), and Methodist (5). Additional denominations listed by fewer than five participants included Christian, Episcopal, Federated Christian, Jehovah's Witness, Pentecostal, Southern Baptist, Unitarian, and United Church of Christ. Four participants listed "None" as their denomination, and 20 left this item blank.

Measures and Procedure. Dimensions of affect were represented by five adjectives, and personality traits were represented by four adjectives[4], listed in Appendix A. Approximately one half of the participants completed the adjective ratings in alphabetical order, and the other half completed the task in reverse-alphabetical order. Participants first were asked to describe an important religious experience that they have had. Then, in the context of that experience, participants rated how well each adjective described their feelings during that experience, and how important the word is for describing religious experiences.

Results and Discussion

Adjectives representing each affect or personality dimension were averaged, yielding scores with theoretical limits of 1 (Does not-) and 7 (Does describe my feelings during the religious experience). For questions regarding how frequently participants experience the word in their religious experiences, responses were anchored by Never and Always. Separate analyses were conducted for affect and personality trait terms.

Affect. A 2 (Intensity: High vs. Low) X 2 (Valence: Positive vs. Negative) X 2 (Congruency: Importance vs. Frequency) within-subjects ANOVA was conducted on participants' ratings of the affect terms. Results revealed several significant effects, including main effects for Intensity, $F(1, 92) = 4.17$, $p < .05$, $MSE = 0.60$, partial $\omega^2 = .02$, and for Valence, $F(1, 92) = 212.09$, $p < .001$ $MSE = 3.66$, partial $\omega^2 = .53$. These main effects were qualified by a significant Intensity X Valence interaction, $F(1, 92) = 37.41$, $p < .001$, $MSE = .45$, partial $\omega^2 = .19$, and a Congruency X Valence interaction, $F(1, 92) = 24.83$, $p < .001$ $MSE = .59$, partial $\omega^2 = .20$. Finally, these effects were moderated by a significant interaction involving all three factors, $F(1, 92) = 5.14$, $p < .05$ $MSE = .18$, partial $\omega^2 = .02$.

In order to understand the nature of the interaction, two separate 2 (Intensity) X 2 (Congruence) ANOVAs were conducted, one for positive affect, and one for negative affect. For positive affect, there was a significant main effect for Congruence, $F(1, 92) = 29.69$, $p < .001$, $MSE = .51$, partial $\omega^2 = .13$, with participants rating these adjectives as being greater in importance ($M = 4.4$) than they were experienced ($M = 4.0$). A second main effect also was found for Intensity, $F(1, 92) = 23.11$, $p < .001$ $MSE = .70$, partial $\omega^2 = .11$. Low intensity affect ($M = 4.0$) was rated as being less relevant to describing religious experiences than was high intensity affect ($M = 4.4$).

For negative affect, only a significant effect for Intensity was found, $F(1, 92) = 9.27$, $p < .01$ $MSE = .35$, partial $\omega^2 = .04$. Although it might be expected that at least one of the two-way ANOVAs would show a significant effect, this is not necessarily true when, as occurred in this data set, the pattern of means are different in the two two-way interactions (Keppel, 1982). Adjectives reflecting high intensity negative affect ($M = 2.2$) were considered more relevant to describing religious experience than were low intensity negative affect adjectives ($M = 2.0$). Of course, these results should be interpreted in the context of a large effect for valence, in which positive valence was more descriptive of religious experience than was negative valence.

Personality Adjectives. There were two main questions of interest regarding the personality adjectives. First, did the general pattern found in Nielsen (2000) hold true in this sample? If so, do differences exist in the ratings of how important the adjectives are, and how frequently the adjectives actually are experienced? If the pattern of results found in earlier research holds true, it would be expected that the adjectives representing low neuroticism, high agreeableness, and high conscientiousness would receive the highest ratings. In order to confirm this pattern of results, data first were analyzed using a 5 (Trait: Neuroticism vs. Openness vs. Extraversion vs. Agreeableness vs. Conscientiousness) X 2 (Pole: High vs. Low) repeated measures ANOVA. The Trait X Pole interaction, $F(4, 368) = 103.36$, $p < .001$, $MSE = 2.52$, partial $\omega^2 = .52$, is quite similar to that found in Nielsen (2000). Adjectives corresponding to the low pole of neuroticism ($M = 4.4$), and the high poles of agreeableness ($M = 4.5$) and conscientiousness ($M = 3.7$), received the highest endorsement by participants.

TABLE 1 Congruence of Frequency and Importance of Trait Adjectives, by Pole, among U.S. Christians

	Congruence			
Pole and Trait	Importance	Frequency	*F* value	*MS* Error
High Pole				
Neuroticism	2.50	2.64	2.29	0.40
Openness	3.07	2.84	7.58*	0.32
Extraversion	3.68	3.39	10.47*	0.36
Agreeableness	4.83	4.21	34.56**	0.51
Conscientious	3.88	3.58	11.51**	0.37
Low Pole				
Neuroticism	4.60	4.10	22.69**	0.50
Openness	3.03	3.10	0.48	0.50
Extraversion	3.43	3.50	0.65	0.35
Agreeableness	2.06	2.08	0.05	0.32
Conscientious	1.97	1.90	0.40	0.48

Note. Higher ratings reflect greater endorsement of the adjectives for describing religious experiences. To protect against inflated family-wise error, a modified Bonferroni procedure (Keppel, 1982) was used, yielding alpha = .02.
$df = (1, 92)$. * $p < .02$ ** $p < .001$

The second question of interest concerned the extent to which participants indicated that they experienced each of these traits. This was tested with a set of planned comparisons. For each pole of each trait, the Importance and Frequency ratings for the adjectives were compared. This was done in order to investigate possible discrepancies between the adjectives that people believe are important in describing religious experiences, but which are not actually achieved or felt. The means corresponding to these contrasts appear in Table 1, along with the results of the planned comparisons. Ratings of Importance and Frequency were incongruent on the High Pole of all traits except for neuroticism. For neuroticism, importance and frequency ratings are incongruent only on the low pole.[5]

The overall pattern described in Nielsen (2000) was replicated in this sample, with participants describing religious experiences in positive terms. Furthermore, the positive nature of religious experiences is reinforced by the lack of congruence between the ideal and the actual experience, as ideal experiences receive more desirable ratings than do the actual experiences. The large effect for valence, noted earlier, is consistent with this conclusion. Finally, participants indicated that high-intensity affect is more characteristic of religious experiences than is low intensity affect, whether positively- or negatively-valenced.

STUDY 2

The research described in MacDonald (2000), Nielsen (2000), and Study 1 all utilized samples made up largely of North American Christians. In order to assess the generalizability of the basic findings of those studies, Study 2 uses a sample of Malaysian Muslims. Relatively little research has been conducted comparing Muslims with members of other religions. Of that which has been done, perhaps the most common theme is the difference between the two religions in terms of deference and independence. Islamic emphasis on religious expertise has been offered to explain high degrees of deference and order needs found among Muslims. This pattern has been observed, for example, in comparisons of Muslims with Hindus living in India (Ojha & Jha, 1991), and Muslims and Christians living in Nigeria (Asonibare, 1986).

In a related vein, Burns and Brady (1992) found that U.S. students scored higher on need for uniqueness than did Malaysian students. It is conceivable that at least part of this effect is attributable to Malaysia's adoption of an official program of Islamization, intended by officials to improve the country (Muzzafer, 1986).

Similar patterns can be found in epidemiological research, which demonstrates that Islamic injunctions against suicide apparently are borne out in suicide rates. Kok (1988) compared Singapore Confucian, Hindu, and Muslim women's suicide rates, and found a pattern consistent with religious beliefs; Muslims were less likely to have committed suicide than were Confucians, while Hindus had the highest suicide rate.

Despite these differences between religious groups, research gives us some reason to expect a degree of similarity in some religious or moral perspectives in people from different religions and cultures. Wolfe and Mourribi (1985) found that Lebanese Christians and Muslims exhibited relatively similar patterns on the Rokeach Value Survey, with honesty, responsibility, freedom and peace rated highly by both groups. At a broader level of analysis, Moon (1986) demonstrated that the basic factor structure and psychometric properties of the Defining Issues Test was consistent in 20 cross-cultural studies, although there is some evidence that the rates of moral development may vary across cultures.

Given that there appears to be several important cross-cultural similarities in general religious values held among Christians and Muslims (Richards & Bergin, 2000; Wolfe & Mourribi, 1985), we expected to find that U.S. Christians and Malaysian Muslims would describe religious experiences using essentially the same kinds of terms. That is, we expected relatively low levels of neuroticism and high levels of agreeableness and conscientiousness to characterize religious experiences.

A second facet of this study concerns the possible distinction between religiosity and spirituality. Some students of religion recommend that scientists focus on spirituality because it may connote less of an institutional affiliation than does religiosity. Alternatively, opponents of this view cite the long use of "religion" to denote many different types of experiences, and not merely institutional ones (Pargament, 1997). In order to address this possible distinction, in Study 2 we asked some participants to describe an important religious experience, while others described an important spiritual experience. Given that other people are part of the religious setting, we expected higher levels of extraversion to characterize those experiences than in the spiritual condition, which does not necessarily involve institutional involvement or support for the experience.

Method

Participants were 62 Muslim students (33 female, 29 male; *M* age = 20.5) living in Malaysia. The participants were enrolled in an English-language program after having demonstrated proficiency in English by successful completion of the TOEFL. After providing demographic data (sex, age, religious affiliation) half of them were asked to "take a few moments to think of the most important religious experience you have ever had," and then they gave a written description of that experience. The other half of the sample was asked to describe their most important spiritual experience. Following the written description, participants described how they felt at the time of the experience by using the 80 7-point semantic differentials listed in McCrae and Costa's (1985, 1987) five-factor taxonomy of personality. One half of the semantic differentials for each trait were reversed in order to control for possible response biases.

Results and Discussion

Descriptions of the experiences first were examined for the extent to which they focused on institutional religious events, such as prayer or Ramadan. Descriptions were rated independently by two assistants, achieving an intraclass correlation coefficient of .92, $p < .05$; discrepancies were resolved by discussion. Five participants omitted the written description, and were excluded from this analysis. Results revealed no difference in participants' use of institutional religious events between the two religious/spiritual conditions, χ^2 (1, $N = 57$) < 1.0, *n.s.*

Ratings of the person's feelings during the experiences were analyzed next. The semantic differential ratings of adjectives that comprise the respective traits were averaged so that a higher score corresponded to the high pole of that trait. Results were analyzed with a 2 (Experience: Religious vs. Spiritual) X 5 (Trait: Neuroticism vs. Extraversion vs. Openness vs. Agreeableness vs. Conscientiousness) mixed-model ANOVA, with trait serving as a within-subjects factor. Results showed a significant effect for Trait, F (4, 240) = 52.30, $p < .001$, MSE = .70, partial ω^2 = .45, and a significant Experience X Trait interaction, F(4, 240) = 4.17, $p < .01$, MSE = .70, partial ω^2 = .17. Planned comparisons revealed significant differences in descriptions of religious and spiritual experiences on the neuroticism, extraversion, and openness dimensions, all Fs (1, 60) > 4.29, all ps < .05, all MSEs > .27, all partial ω^2s > .05. Neuroticism was rated lower in the religious (M = 3.0) than in the spiritual (M = 3.6) condition. Openness was greater in the religious (M = 4.7) than in the spiritual (M = 4.4) condition. Extraversion also was rated higher in the religious (M = 4.8) than in the spiritual (M = 4.4) condition. Ratings of agreeableness approached significance, F (1, 60) = 3.13, p = .08, with means of 5.5 and 5.1 for the religious and spiritual conditions, respectively.

Taken as a whole these results indicate a general similarity in participants' descriptions in the two studies. Following a pattern consistent with earlier research (Spilka, Ladd, McIntosh & Milmoe, 1996), participants emphasized the positive elements of their religious experiences. Nevertheless, the participants in this study described religious experiences more positively than they described spiritual experiences. People described feeling less neurotic, and more extraverted and open, in the context of religious experiences than in spiritual experiences.

GENERAL DISCUSSION

A problem in the study of religious experiences is that those experiences are ineffable, defying description. For this reason, participants in these studies were asked to use a standard set of adjectives to describe how they felt during the experience, rather than to describe the experience itself. Adjectives reflecting high levels of emotional stability and agreeableness were most consistently associated with religious experience. To a lesser extent, adjectives reflecting higher levels of conscientiousness and extraversion also were useful in characterizing how participants felt during religious experiences. A similar pattern described spiritual experiences. Adjectives associated with positive affect also described participants' religious experiences.

Participants' ratings in these studies indicate the positive nature of religious experiences in this sample. This might be expected in the U.S. sample, which by virtue of its being a snowball sample might be biased in favor of religious experiences. The positive nature of religious experiences, however, also was evident in the Malaysian sample, which consisted of all students in two university classes. The results of this research suggest that people who have had religious experiences rate them very favorably and, as demonstrated in Study 1, they would like to have those experiences more frequently than they actually occur.

It would be desirable in future research to ascertain the distinction, if any, that people make between "religious" and "spiritual" experiences. In this study, participants described religious and spiritual experiences similarly, although religious experiences were described in somewhat more positive terms than were spiritual experiences. This was true despite the fact that the descriptions of those experiences did not differ in their inclusion of institutional events. The generalizability of this finding should be tested, so that we can understand the extent to which the pattern is replicated in other cultures. Perhaps the Malaysian government's official promotion of Islam may have influenced this pattern of results.

Future research also should seek to replicate these results in samples of other religions. In particular, research concerning religious experience in non-monotheistic traditions is particularly lacking; this method might be profitably applied to such religions. Because of the difficulty of creating questionnaire scales that apply equally well to diverse religious traditions, adjectives such as those used in these studies might enable researchers to begin to compare people's experiences in spite of the vast doctrinal differences that exist.

REFERENCES

Allport, G. W., & Odbert, H. S. (1936). Trait names: A psycho-lexical study. *Psychological Monographs, 47* (1, Whole No. 211), 1-171.

Asonibare, B. (1986). Personal control and religious differences in selected secondary school students in Ilorin. *Nigerian Journal of Guidance and Counseling,* 2, 95-102.

Burns, D. J., & Brady, J. (1992). A cross-cultural comparison of the need for uniqueness in Malaysia and the United States. *Journal of Social Psychology, 132*, 487-495.

Crapps, R. W. (1986). *An introduction to psychology of religion*. Macon, GA: Mercer Univ. '

Gallup, G., & Jones, S. (1989). *One hundred questions and answers: Religion in America*. Princeton, NJ: Princeton Religion Research Center.

Henry, G. T. (1990). *Practical sampling*. Newbury Park, CA: Sage.

Hill, P., & Hood, R. W. Jr. (Eds.) (1999). *Measures of religiosity*. Birmingham, AL: Religious Education Press.

Hood, R. W. Jr. (1975). The construction and preliminary validation of a measure of reported mystical experience. *The Journal for the Scientific Study of Religion, 14*, 29-41.

Keppel, G. (1982). *Design and analysis* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Kok, L. P. (1988). Race, religion and female suicide attempters in Singapore. *Social Psychiatry and Psychiatric Epidemiology, 23*, 236-239.

MacDonald, D. A. (2000). Spirituality: Description, measurement, and relation to the five factor model of personality. *Journal of Personality, 68*, 153-197.

McCrae, R. R., & Costa, P. T. Jr., (1985). Updating Norman's "adequacy taxonomy:" Intelligence and personality dimensions in natural language and in questionnaires. *Journal of Personality and Social Psychology, 49*, 710-721.

McCrae, R. R., & Costa, P. T. Jr., (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*, 81-90.

Moon, Y. L. (1986). A review of cross-cultural studies on moral judgment development using the Defining Issues Test. *Behavior Science Research, 20*, 147-177.

Muzaffer, C. (1986). Malaysia: Islamic resurgence and the question of development. *Sojourn, 1*, 57-75.

Nielsen, M. E. (2000). Descriptions of religious experience using trait and affect adjectives. *Psychological Reports, 86*, 308-310.

Ojha, H., & Jha, C. N. (1991). Personality pattern of Hindu and Muslim college students. *Psychologia, 34*, 248-253.

Pargament, K. (1997). The psychology of religion and spirituality? Yes and no. *Psychology of Religion Newsletter of APA Division 36, 22*(3), 1-9.

Preston, C. A., & Viney, L. L. (1986). Construing God: An exploration of the relationships between reported interaction with God and concurrent emotional experience. *Journal of Psychology and Theology, 14*, 319-329.

Richards, P. S., & Bergin, A. E. (2000). Religious diversity and psychotherapy: Conclusions, recommendations, and future directions. In Richards & Bergin (Eds.) *Handbook of psychotherapy and religious diversity* (pp. 469-489). Washington, DC: American Psychological Association.

Rogers, C. (1976). Carl Rogers [Interview]. In R. I. Evans (Ed.), *The making of psychology* (pp. 213-226). New York: Knopf.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 36*, 1161-1178.

Russell, J. A. (1983). Pancultural aspects of the human conceptual organization of emotions. *Journal of Personality and Social Psychology, 45*, 1281-1288.

Spilka, B., Brown, G. A., & Cassidy, S. A. (1992). The structure of religious mystical experience in relation to pre- and postexperience lifestyles. *The International Journal for the Psychology of Religion, 2*, 241-258.

Thomas, L. E., & Cooper, R. E. (1978). Measurement and incidence of mystical experience: An exploratory study. *Journal for the Scientific Study of Religion, 17*, 433-437.

Thomas, L. E., & Cooper, R. E. (1980). Incidence and psychological correlates of intense spiritual experiences. *Journal of Transpersonal Psychology, 12*, 75-85.

Wolfe, G., & Mouribbi, A. (1985). A comparison of the value systems of Lebanese Christian and Muslim men and women. *Journal of Social Psychology, 125*, 781-782.

Notes

1. This dimension has been present in some studies of emotion and religion, e.g., Preston and Viney's (1986) study of perceived relationships with God, and Crapps's (1986) psychoanalytic view of emotion and religion.

2. Although it may be thought that there would be denominational differences, for example, between "mainline" and "Pentecostal" denominations, the present data would not lend themselves to a test of this hypothesis.

3. Informed consent was given by all participants in these studies, and by the guardian of the one minor who participated in this study.

4. Because of a clerical error, the low pole of the openness trait was represented by only three adjectives.

5. Separate analysis of the responses from the 25 participants who had not attended church services during the preceding month showed the same pattern of results, although the non-churchgoing participant ratings of the frequency with which they have the experiences was lower than that for the larger sample.

APPENDIX A
ADJECTIVES USED IN STUDY 1

Personality Terms

Nigh Neuroticism: High-strung, Insecure, Nervous, Worrying
Low Neuroticism: At-ease, Calm, Relaxed, Secure
High Extraversion: Affectionate, Fun-loving, Sociable, Talkative
Low Extraversion: Quiet, Reserved, Retiring, Sober
High Openness: Creative, Daring, Imaginative, Original
Low Openness: Conventional, Down-to-earth, Unadventurous
High Agreeableness: Forgiving, Good-natured, Soft-hearted, Sympathetic
Low Agreeableness: Callous, Irritable, Ruthless, Vengeful
High Conscientiousness: Careful, Conscientious, Reliable, Well-organized
Low Conscientiousness: Careless, Disorganized, Negligent, Undependable

Affect Terms

High Intensity Positive: Enthusiastic, Excited, Joyful, Stimulated, Zestful
High Intensity Negative: Afraid, Aggravated, Distressed, Frustrated, Miserable
Low Intensity Positive: At-ease, Calm, Content, Relaxed, Satisfied
Low Intensity Negative: Bored, Discontent, Dejected, Gloomy, Tense

The Best Laid Plans of Mice and Men: The Role of Decision Confidence in Outcome Success

Prasad Bingi & David L. Turnipseed
Indiana-Purdue University

George M. Kasper
Virginia Commonwealth University .

This paper develops an interactive model depicting the role of decision confidence in outcome success by integrating relevant knowledge from decision making, psychology, strategic management, and organizational behavior literatures. Analyzing the hypothesized relationships using linear structural modeling produced results indicating that decision confidence has a positive impact on implementation effort, and implementation effort directly effects outcome success. These findings elucidate the importance of confidence, which may be as important as quality in determining the success of a plan.

Research has consistently shown that explicit and relatively challenging goals produce higher levels of performance than generalized and easy goals (e.g., Locke & Latham, 1990; Locke, Shaw, Saari, & Latham, 1981). However, goals may not produce sustained effort in absence of performance knowledge or feedback (Bandura & Cervone, 1983). The literature indicates that goals are more likely to motivate toward higher performance when individuals are committed to them. Research in strategic planning suggests that these task-goals attributes are consistent for organizations as well as individuals (Boal & Bryson, 1987).

Dean and Sharfman (1996) provided evidence that decision making processes are related to outcome success. Outcome success is a measure of effort, but decision quality is a measure of process. In the strategic management literature, decision quality is often used synonymously with the quality of a plan or, as stated by Koontz (1958), "Planning is deciding..." (p. 48). Thus decision confidence may be equated with confidence in a plan. Plans are often more expansive then decisions, and include specified means-ends relationships as well as ends to be achieved. In either case, a good plan is high in decision quality and a poor plan is low in decision quality. Plans influence outcomes, but good plans do not

---

Author info: Correspondence should be sent to: David Turnipseed, Business & Mgmt. Sciences, Indiana-Purdue University, Fort Wayne, IN 46805-1499.

North American Journal of Psychology, 2001, Vol. 3, No. 1, 91-108.
© NAJP.

necessarily produce good outcomes, and poor plans do not always lead to poor results. Outcome success is only a partial measure of the quality of a decision. It is the connection between the decision making process and its outcome which is addressed in this study.

We propose that decision confidence, or belief in the quality of a decision, is an indispensable link between planning and attaining optimal results. Vroom and Yetton (1973) argued that when subordinates are required to use their initiative in implementing a decision, acceptance and, by implication, commitment to the decision is crucial. Decision confidence is distinct from self-confidence, self-efficacy, outcome expectancy, trust in a decision aid, or other types of confidence. Decision confidence affects implementation effort, outcome success (Russo & Schoemaker, 1992), and the amount of implementation effort devoted to a course of action in pursuit of an outcome (Stevick, Martin, & Showalter, 1991). Confidence also plays a key role in the selection stage of decision making, as an individual chooses among viable alternative decisions (Kasper, 1996).

The role of decision confidence in outcome success "... is critical when it comes to implementation..." (McIntyre, 1982, p. 17), as confidence provides the drive to implement decisions. An individual's confidence in a decision influences determination and the persistence to pursue a given course of action (Neale & Bazerman, 1985). High confidence in a poor decision or low confidence in a good decision may both be undesirable. Confidence greater than warranted by the quality of the decision can result in a wasteful over-commitment of resources as well as failure to recognize the inherent situational risks and to take risk-commensurate precautions. Overconfidence can also lead to an escalation of commitment and persistence in implementing a decision despite clear evidence of its ultimate failure (Staw, 1976; 1981). Conversely, decision confidence less than warranted by the quality of a decision can mean that too few resources and efforts are devoted to implementation. Other possibilities include foregone opportunities, premature abandonment of a good decision, or implementation efforts that fail to realize the full potential of the decision.

The motivational benefits of decision confidence are linked to successful implementation (Hogarth, 1980; Russo & Schoemaker, 1992; Trull, 1966). Hogarth (1980) contended that for creatures needing to master their environment to survive, decision confidence characterized by over reaction is far less costly than confidence characterized by inaction. As Weick (1990) stated, when one is lost in the desert, any map is better than no map. In business, for example, it is evident that planning for the future is more successful than simply reacting to the actions of others. Unfortunately, the same confidence that motivates in the case of good decisions can be disastrous with bad decisions, as any manager who has 'thrown good money after bad' will attest (Conlon & Wolf, 1980; Fox & Staw, 1979; Staw, 1973;1981).

Goal assigners are likely to be authority figures, and thus goals may fall into the zone of indifference (Barnard, 1938) which one accepts as part of the psychological contract with the organization (Rousseau, 1996). Individuals with the authority to assign goals are typically viewed as having the right, if not the expertise, to determine those goals. The act of assignment directly implies that the subordinates can attain the goal. We believe that confidence in a goal is a significant determinant of individuals' acceptance and commitment to the goal, which subsequently influences the probability of a successful outcome.

Sniezek (1992) and McIntyre (1982) proposed that decision confidence mediates efforts devoted to pursuing a course of action, and those efforts affect the outcome success. Confidence appears to be critical in determining commitment, implementation effort, and ultimately, outcome success. Consequently good decisions alone may be insufficient to produce desired outcomes. Prior research has focused on realism in decision confidence, largely ignoring its motivational benefits and its potential impact on implementation effort. Studies have concentrated on decision quality, neglecting decision confidence and its role in outcome success. No single study has considered the relationship between decision confidence, implementation effort and outcome success. Only anecdotal evidence exists regarding the influence of decision confidence on implementation efforts and outcome success. We believe, consistent with Sniezek & Henry (1989), that confidence in the plan may be as important as quality to outcome success.

The purpose of this study is to investigate the effect of decision confidence on outcome success. A decision confidence model is developed based on decision making, strategic management, motivational psychology, and organizational behavior literatures, and critical aspects of the model are empirically assessed with structural equation analysis.

Integrating the findings reported in the literature regarding relationships between decision confidence, implementation effort, and outcome success, produced the model proposed in Figure 1. For simplicity, the model assumes decision making and implementation by an individual; however, the model can easily be extended to include group situations. The research issues are divided into primary and secondary hypotheses. Hypotheses 1-3 are the primary hypotheses; the other 7 hypotheses are secondary. Each hypothesis is designated by a number on the paths of the model in Figure 1.

Figure 1 proposes that outcome success is directly affected by decision quality and implementation effort. Through implementation effort, outcome success is indirectly affected by decision confidence, goal commitment, self-efficacy, goal-setting behavior, and feedback. A decision-maker's role in the generation of computer output is also believed to affect outcome success through decision confidence. According to this model, self-efficacy influences both goal-setting behavior ("self-set" goal in Figure 1) and goal commitment, thereby indirectly affecting outcome success.

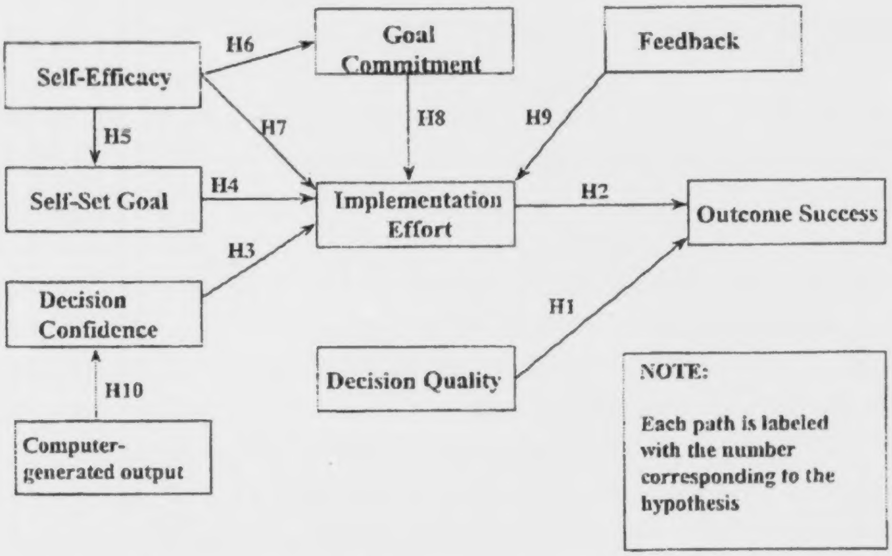

Self-Efficacy
H6
Goal Commitment
Feedback
H5
H7
H8
H9
Self-Set Goal
H4
Implementation Effort
H2
Outcome Success
H3
Decision Confidence
H1
H10
Decision Quality
Computer-generated output
NOTE:
Each path is labeled with the number corresponding to the hypothesis

Figure 1. The Research Model with Hypotheses Marked on the Paths

Outcome success is dependent upon decision quality, but they are equivalent in the absence of exogenous variables. Because the real world contains numerous uncontrollable variables, decision quality is rarely, if ever, equivalent to outcome success. Conventional wisdom and the limited evidence indicates a relationship between decision quality and outcome success (e.g., Beyth-Marom, Dekel, Gombo, & Shaked, 1985), and link 1 of Figure 1 proposes a positive relationship between decision quality and outcome success. Thus the following hypothesis:

H_1: *Decision quality has a positive impact on outcome success.*

Link 2 considers the effect of implementation effort on outcome success. The strategic management literature firmly establishes that without proper implementation, decisions are ineffective (e.g., Cravens, 1991; Digman, 1990; Thompson & Strickland, 1999), which suggests the following hypothesis:

H_2: *Implementation effort has a positive effect on outcome success.*

Decision confidence has been argued to influence implementation effort as illustrated by link 3 (McIntyre, 1982). We hypothesize that the level of implementation effort depends upon the degree of decision confidence:

H_3: *Decision confidence has a positive effect on implementation effort.*

A relationship between decision confidence and decision quality is not shown in Figure 1 because both decision confidence and decision quality are treated as independent variables. Hence the model does not include a link between the decision quality and confidence constructs.

The secondary hypotheses involve implementation effort, self-efficacy, goal-setting behavior, and feedback. Decision-making involves goal setting (Beach, 1985), and according to goal-setting theory, goals have some degree of motivational value and they influence implementation efforts (Locke & Latham, 1990). Goals direct and energize actions as indicated by link 4, and thus the hypothesis:

H_4: *Goal setting is positively related to implementation effort.*

The personal conviction that the plans required to achieve a desired outcome can be successfully implemented is also important to outcome success. This self-efficacy regulates one's effort and persistence in goal directed activities (Bandura, 1977), and influences goal commitment, implementation effort, and outcome success (Frayne & Latham, 1987; Katzel & Thompson, 1990; Locke, Frederick, Lee, & Bobko, 1984; Locke & Latham, 1990). Goal-setting and self-efficacy theories predict that self-efficacy affects goal setting (link 5), goal commitment (link 6) (Locke & Latham, 1990), and implementation effort (link 7) (Bandura, 1977). These self-efficacy relationships suggest the following hypotheses:

H_5: *Self-efficacy has a positive effect on goal setting.*

H_6: *Self-efficacy has a positive effect on goal commitment.*

H_7: *Self-efficacy has a positive effect on implementation effort.*

Goal commitment is linked to outcome success (Locke & Latham, 1990). Transferring a decision into implementation effort requires an appropriate level of commitment (Wernham, 1987; Milliken & Vollrath, 1991). There is a relationship between goal commitment and the willingness to exert effort: all other variables being constant, outcome success depends upon goal commitment and effort (Guth & MacMillan, 1986), thus Hypothesis 8:

H_8: *Goal commitment has a positive effect on implementation effort.*

Prior research has shown that goals combined with feedback influence outcome success (Locke & Latham, 1990). Feedback provides a cognitive comparison between goals and the individual's progress, thus regulating via the self-efficacy mechanism, the implementation efforts required to achieve a desired outcome (Earley, 1988; Earley, Northcraft, Lee & Lituchy, 1990).

Goals have been observed to produce sustained performance effort when combined with feedback of progress (Bandura & Cervone, 1983). Hence we hypothesize:

H_9: *Feedback has a positive effect on implementation effort.*

Considering the impact of computer-generated plans on decision confidence and outcome success is particularly relevant as decisions are increasingly made based on computer-generated information. The limited research suggests that computer-generated material influences decision confidence (Freeman, 1984; Huff, 1986). Likewise, self-generated output, with or without the help of a computer-based system, has been shown to affect confidence in the output (Northcraft, & Earley, 1989). Whether and how changes in decision confidence resulting from the use of computer-generated plans influence outcome success is unknown. The sparse research suggests that self-generated computer output may have more influence on decision confidence than output generated by others, providing the hypothesis:

H_{10}: *Self-generated computer produced plans will have a greater effect on decision confidence than other-generated output.*

METHOD

To test the hypotheses, decision confidence, decision quality, computer output, and feedback were manipulated. Decision confidence and quality were manipulated to produce high and low levels. Self-generated computer-based plans were compared with other-generated plans, and subsets of the participants were provided with periodic feedback on their performance progress. Participants were randomly assigned to one of the manipulation groups or to a non-manipulated control group.

To capture the desired goal attainment data, decision confidence and quality were manipulated and the variance in implementation effort and outcome success was measured. We selected an experimental "decision" that was neither extremely easy nor inordinately difficult for the participants, and one that would motivate students to participate in the study. Other criteria included the relevancy of the study to students as well as an easily understandable protocol. Consistent with these requirements, participants were given the task of improving their reading skills. Participants were given an individualized plan developed by a reading specialist, designed to help them reach a desired reading skill over the five-week study period, which is typical of traditional methods used to improve reading.

The primary dependent variable of interest was outcome success. Other endogenous variables were implementation effort, goal commitment, and goal-setting behavior. Consistent with Bonoma and Clark's (1988) recommendation, both objective and subjective measures of implementation effort and outcome success variables were recorded. The objective measures of outcome success and goal-setting behavior were based on the Nelson-Denny Reading Test (NDRT) (1981). The NDRT score, which is a linear combination of vocabulary and comprehension scores, is a well known and popular measure of reading skill (Brown, Bennett & Hanna, 1981).

Goal setting behavior was assessed at the beginning of the experiment. The NDRT was administered, scored, and returned to the participants. Based on the NDRT total scores, five possible goals were identified, and following Locke and Latham (1990), the goals were defined so that 10%, 25%, 50%, 75% and 90% of the participants could reach the goal. This resulted in goals of 150, 130, 115, 95 and 75 as NDRT total scores, respectively. Each participant chose one of these as his/her goal. At the end of the study, the NDRT was re-administered. The objective measure of outcome success was calculated as the ratio of final performance (post NDRT total score) to the target performance (goal). The subjective measure of success was the participants' satisfaction with their reading skill scores at the end of the study period.

Implementation effort was also measured with both objective and subjective measures. Objective measures included the amount of time the participants spent in formal practice lessons conducted by the reading specialists, formal training sessions, and volunteer help sessions. Participants were not aware that these values were being recorded. These measures were averaged to obtain objective measures of implementation effort.

The subjective measure of implementation effort was the instrument developed by Earley, Wojnaroski and Prest (1987). Using the Likert scale instrument, each subject estimated the amount of time he/she spent outside of the formal settings completing each lesson. Earley, Wojnaroski and Prest (1987) reported an alpha of .73 for this questionnaire. Goal commitment was assessed using the instrument developed by Hollenbeck, William and Klein (1989), which measures commitment on a five-point Likert scale, and formed one factor with an alpha of .88.

The independent variables were: (1) decision quality, which was the quality of the plan given to the subject to achieve his/her goal; (2) decision confidence, the subject's belief in the quality of the plan; (3) self- or other-generated computer output (plans); and (4) the presence or absence of feedback. Self-efficacy was a measured but not manipulated independent variable.

Individuals in real-world situations are typically given a decision in the form of a plan that someone else has developed, and they accept that decision with varying degrees of confidence in its quality. In this study, a professional reading specialist developed a high decision quality (effective) or a low decision quality (ineffective) reading plan for each subject at random, based the subject's performance on the NDRT, the goals set, and study habits reported. The study habits had been collected from a biographical questionnaire administered at the beginning of the experiment. These plans were double-blind rated by two other independent professional reading specialists employed for the study. Using a standard instrument, the three reading specialists reported an inter-rater reliability of .94 (only 14 observations out of 253 were in disagreement) for the decision quality of the reading plans given to the participants.

Decision confidence was manipulated by the reading improvement plans. Participants were told that their reading plans would be generated by an "expert system" based on their demographic information, NDRT score, and stated goals. High and low confidence reading plans were generated. The high decision confidence plan described the expert system used to generate the plan as extensively tested and highly reliable and accurate, with 95% the plans generated by the system being effective. The low confidence plan described the expert system as developmental, not extensively tested, and although the reliability and accuracy of the generated plans were unknown, only 5% of the plans generated by the system so far had been effective. To assess the effectiveness of this manipulation, participants completed a Decision Confidence in the Quality of the Plan instrument.

Feedback was also manipulated between subject groups with one group receiving feedback in the form of the answer key to practice lessons which provided instantaneous assessment of their goal-performance discrepancy. Participants in the no-feedback group did not have an opportunity to assess their interim performance. Precautions were taken to prevent or reduce the feedback group from contaminating the no-feedback group. For example, feedback participants were given the answer key for the current practice lesson they were completing and only for the material related to their reading plan. Therefore, even if the answer key became available to those in the no-feedback group, it was highly unlikely that it would be useful to those participants on their practice lessons.

Self-efficacy was conceptualized in terms of magnitude and strength, consistent with Bandura (1977). These were measured based on Locke and Latham's (1990) recommendation: magnitude was measured by asking the participants to indicate which of the five goals (i.e., improving their NDRT by 150, 130, 115, 95, or 75 total points over the course of the study) they felt that they could achieve. Strength was assessed by asking the participants to indicate on a scale of 1 to 10 (1 = no confidence, 10 =

highest confidence) how confident they were in their ability to achieve each of the five goals.

Participants

Students enrolled in a junior-level, required business course were solicited to participate in the study. Participation was voluntary, and students who participated earned class credit. Students choosing not to participate were given other projects to earn the same amount of credit. Data from 279 participants was used in the analysis. Approximately 40% of these par- ticipants were female, and about 50% were juniors between the ages of 18-22. Fifteen of the participants were used as a control group, and data from the remaining 253 participants was used to investigate the hypotheses.

Procedure

Prior to the experiment, a biographic questionnaire was administered to determine the participants' major area of study, GPA, and, reading habits. Participants were randomly assigned to one of the test groups defined by the manipulations. The NDRT was administered and scored, and the results (vocabulary, comprehension, and total scores) provided to the participants. Participants completed the self-efficiency instrument by indicating which of the five goals (i.e., improving their NDRT by 150, 130, 115, 95, or 75 total points over the course of the study) they felt they could achieve (yes or no) over the during the five-week period, and how confident they were in their ability to achieve each of the five goals (1 = no confidence, 10 = highest confidence). Participants selected one of the five goals as their reading score improvement goal. Often, this was not the goal to which the subject assigned the highest confidence, which is entirely consistent with normal human behavior, as many individuals are not high achievers in various endeavors although they may have confidence that better performance is possible. Finally, goal commitment was measured using the instrument developed by Hollenbeck, William, & Klein (1989).

Participants were provided their reading plan by either the "chauffeur" or the self-generated treatment as dictated by the experimental design. Participants in the "chauffeur" treatment were told that their plan had been generated using the same computer program that the other (self-generated) group was using. Accompanying each plan was information about the performance of the fictitious "expert system" used to produce the plan, which was used to manipulate the subject's decision confidence in the plan.

In reality, high and low quality reading plans had previously been developed by the reading specialist, and based on the participant's experimental group, either the high or low quality plan was provided by the system. Participants in the "chauffeur"-generated group were given their reading plans by the researchers and told that the reading plan was generated for them using the same "expert system" that the self-generated group was using. Each reading plan consisted of a regime of training sessions to be attended, due dates for practice lessons, details for help sessions, and grading policies. Participants were required to attend each of the training sessions (set out in their reading plans) which were conducted by a reading specialist. At the end of each training session, practice assignments were given and participants were required to turn in the completed assignments along with an approximate amount of time they spent on the practice lesson. Help sessions were provided by the reading specialist to all participants, and the time spent in the help sessions was recorded and included as one measure of implementation effort. Finally, the questionnaire assessing the efficacy of this manipulation, the Decision Confidence in the Quality of the Plan questionnaire, was administered.

During the course of the experiment, participants in the feedback group were given the answers to their practice lessons. Participants in this group were told that the program needed to be completed by the end of the five-week period, and because the instructor could not grade all the practice lessons in a timely manner, they had been requested to help by grading all their own practice lessons. Both overt and covert inquiries throughout the experiment indicated that there was little or no contamination between the feedback and no-feedback groups.

At the end of the five-week experiment period, the NDRT was re-administered to measure the change in the participants' reading scores. After the NDRT answer sheets were scored and returned, each subject then indicated his or her satisfaction with the performance by completing the Outcome Success Satisfaction instrument.

RESULTS

The efficacy of the manipulations was verified using ANOVA. Structural equation analysis was chosen to investigate the hypotheses because it allows testing of simultaneous equations and whole models, is useful when there are several constructs underlying dependent variables, provides a way to correct measurement error, and violations of homogeneity can be treated (Bagozzi & Yi, 1989).

ANOVA of the pre-and post-treatment NDRT showed that the participants exposed to the manipulations and reading regimes performed significantly better than those in the control group ($F = 6.71$; $p < .01$) who received neither manipulations nor reading regimes. To verify the decision confidence manipulation, a *t*-test was calculated, which indicated that the decision confidence of participants in the high confidence group was greater than that of participants in the low confidence group ($p < .05$).

Factor analysis of the measures of outcome success, implementation effort, and goal commitment were unidimensional for each construct. Items which constructed each of these measures cleanly loaded on three appropriate dimensions, providing support for the validity of the scales. The measures of outcome success combined the objective (the ratio of the subject's actual reading skill score to his or her goal reading skill score) and the subjective assessment of outcome success (the participant's satisfaction with his or her reading skill score at the end of the study period). The measure of implementation effort combined the average of the amount of time the subject spent in formal practice lessons, formal training sessions, and voluntary help sessions, and the subjective measure of implementation effort. The goal commitment instrument was also reduced to a single measure. Reliability estimates (Cronbach's alpha) for each of these measures was greater than .78.

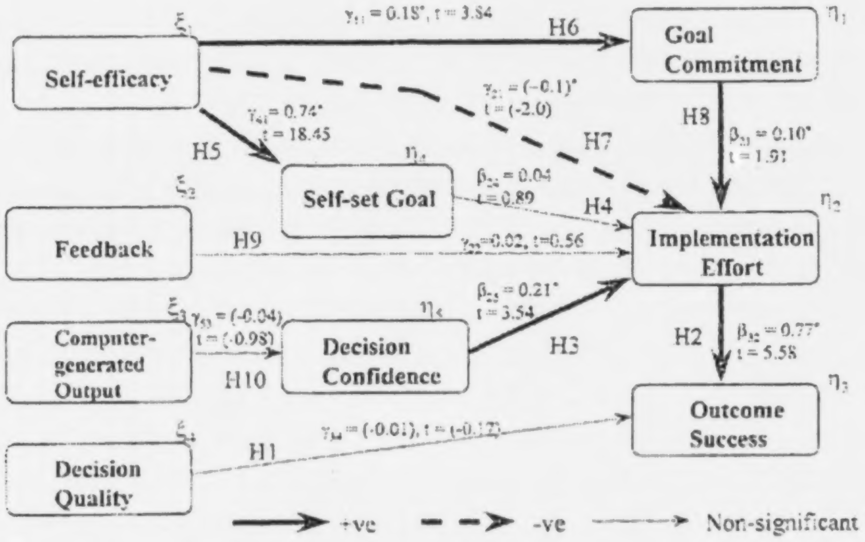

ξ1
Self-efficacy
γ11 = 0.18*, t = 3.84
H6
Goal Commitment
η1
γ21 = (−0.1)*
t = (-2.0)
H7
H8
β21 = 0.10*
t = 1.91
γ41 = 0.74*
t = 18.45
H5
η4
Self-set Goal
β24 = 0.04
t = 0.89
H4
ξ2
Feedback
H9
γ22 = 0.02, t = 0.56
Implementation Effort
η2
ξ3
γ53 = (-0.04)
t = (-0.98)
Computer-generated Output
H10
η5
Decision Confidence
β25 = 0.21*
t = 3.54
H3
H2
β32 = 0.77*
t = 5.58
η3
Outcome Success
ξ4
γ34 = (-0.01), t = (-0.12)
H1
Decision Quality
+ve
-ve
Non-significant

Figure 2. Structural Model with Results of Structural Analysis

---

Goodness-of-fit: $\chi^2 = 274.02$ (d.f. = 215, p = 0.004), GFI = 0.92, AGFI = 0.89, NFI = 0.97, NNFI = 0.97, CFI = 0.98, IFI = 0.98, RMR = 0.057, RMSEA = 0.033

* Significant at 0.05 level; t-values greater than 1.65 are significant at $p \leq 0.05$ (one-tailed test)

Confirmatory factor analysis was conducted on the reduced, factor-analyzed scales. These results showed that all the constructs were unidimensional. The resulting goodness of fit statistics were: $\chi^2_{(198)} = 248.9$ ($p < .008$); goodness-of-fit index (GFI) = .92; adjusted goodness-of-fit index (AGFI) = .90; root mean square residual (RMR) = .045; root mean square of approximation (RMSEA) = .03; normed fit index (NFI) = .92; non-normed fit index (NNFI) = .98; and comparative fit index (CFI) = .98. Of all the pairwise correlations between measurement items, only 4.1% (11) had normalized residuals greater than 2.0, which is less than the 5% expected by chance. Collectively, these measures suggest that the data are consistent with the model in Figure 1.

A structural model was estimated using maximum-likelihood in LISREL VIII (Jöreskog & Sörbom, 1993). The correlation matrix of the sample data was used to test the overall adequacy of the theorized model as well as the significance of the hypothesized relationships. Analysis by the procedures recommended by Anderson and Gerbing (1988) confirmed the discriminant and convergent validity of the instruments, with none of the confidence intervals containing 1.0. The strongest correlation for outcome success was implementation effort ($r = .43$). For implementation effort, the next strongest correlation was decision confidence ($r = .28$). Other robust correlations include self-efficacy and goal-setting behavior ($r = .76$), self-efficacy and goal commitment ($r = .29$), and goal commitment and goal-setting behavior ($r = .25$).

The hypothesized relationships (paths), structural coefficients, *t*-values, and overall goodness of fit indices are provided in Figure 2. Traditional indices ($\chi^2_{(215)} = 274.02$, $p < .01$, GFI = .92, AGFI = .89, NFI = .91, NNFI = .97 and RMSEA = .033), the incremental fit index (IFI = .98), and the comparative fit index (CFI = .98), indicate that our model fits the data. Additionally, the point estimate of RMSEA is well below the .05 level recommended by Browne and Cudeck (1993), confirming the good fit. The Expected Cross Validation Index (ECVI) for the model (1.57) is less than the ECVI for the saturated model (2.19), further supporting our model.

The structural model was also compared to the measurement model (Figure 1) using the sequential chi-square difference test (SCDT) as recommended by Anderson and Gerbing (1988). The chi-square for the measurement model ($\chi^2_{m\ (198)}$) was 248.93, and for the structural model ($\chi^2_{s(215)}$), 274.02. Differences between these chi-squares was not significant ($\chi^2_s - \chi^2_m = \chi^2_d = 25.02$, 17 *df*), indicating that the restrictions placed on the measurement model are reasonable.

Figure 2 also presents the coefficients and t-values for each of the paths in the structural model. The hypothesis, dependent and independent variables, the test result, and summaries of the implications are presented in Table 1.

TABLE 1 Summary of Hypotheses Tests

Hypothesis	Dependent Variable	Test Result	Implication
1	Outcome success	Not Supported	Decision quality had no effect on outcome success.
2	Outcome success	Supported	Implementation effort had a positive effect on outcome success.
3	Implementation Effort	Supported	Decision confidence had a positive effect on implementation effort.
4	Implementation Effort	Not Supported	Goal-setting behavior had no effect on implementation effort.
5	Goal-setting Behavior	Supported	Self-efficacy had a positive effect on goal-setting behavior.
6	Goal Commitment	Supported	Self-efficacy had a positive effect on goal commitment.
7	Implementation Effort	Not Supported	Self-efficacy had a negative effect on implementation effort.
8	Implementation Effort	Supported	Goal commitment had a positive effect on Implementation effort.
9	Implementation Effort	Not Supported	Feedback had no effect on implementation effort.
10	Decision Confidence	Not Supported	Decision confidence was not affected by the type of computer-generated output (self-generated or Chauffeur-generated).

DISCUSSION

The analysis provided support for two of the three primary hypotheses of interest. H_2 was supported, suggesting that implementation effort has a positive effect on outcome success. Likewise, H_3 was supported, sustaining our assertion that decision confidence has a positive effect on the implementation effort. Taken together, the results suggest that decision confidence affects implementation effort, which affects outcome success. Participants with high implementation effort were more likely to achieve their self-set goal and be satisfied with their performance than participants which had expended little implementation effort. Decision quality was also expected to affect outcome success, and although participants with high quality plans performed marginally better than those with low quality plans, the difference was not significant, and H_1 was not supported.

The support of the secondary hypotheses H_5 and H_6 indicate that goal commitment directly affects implementation effort, and self-efficacy directly affects both goal-setting behavior and goal commitment: additionally self-

BEST COPY AVAILABLE

efficacy indirectly affects implementation effort. Hypothesis H_8 was supported, indicating that goal commitment has a positive impact on implementation effort.

H_4 was rejected: our data showed that goal setting behavior had no effect on implementation effort. H_7 was also rejected, with self-efficacy having a negative rather than the hypothesized direct effect on implementation effort. H_9 was not supported, indicating that feedback does not affect implementation effort. Finally, as the data show no evidence that self-generated computer output affected decision confidence more than chauffeur-generated computer output, H_{10} was rejected

Our results support the basic proposition that decision confidence affects implementation effort. Participants in the high decision confidence group expended greater effort implementing their reading plans than participants in the low confidence group. Because decision confidence has been shown to affect implementation effort, researchers should devote increased attention to decision confidence and its antecedents. Practitioners should be aware of the effects of decision confidence on implementation effort, and realize that good planning alone is not sufficient for good outcomes.

Outcome success requires high implementation effort, and high implementation effort requires confidence in the decision. Among the secondary findings, self-efficiency emerged as an important contributor to outcome success. Self-efficacy was positively related to the level of the goal set by the subject as well as his/her level of commitment to that goal. This implies that individuals with high self-efficacy tend to choose harder goals and have increased commitment to those goals. However, self-efficiency had a negative impact on implementation effort. Perhaps participants with little self-efficacy overcompensate by exerting high implementation effort to achieve their goals. Participants with high self-efficiency may feel confident in their ability to achieve a goal and consequently fail to apply themselves and exert the implementation effort needed to execute the plan.

Consistent with Locke, Shaw, Saari, and Latham (1981), participants with high goal commitment also exerted high implementation effort. Goal-setting behavior had no affect on implementation effort, however this unexpected result has been observed previously by Locke and Latham (1990), who referred to several studies which showed no effect of goal-setting behavior on implementation effort. They attributed these results to, among other things, lack of training, or the mismatch between performance criteria and goals or feedback. These shortcomings and inconsistencies may also be responsible for the results in the present study.

Feedback had no influence on implementation effort, which may have been due to the type of feedback employed. Research has shown that increased motivation is possible only when the feedback matches the performance criteria (Earley, Wojnaroski, & Prest, 1990). The feedback provided to our participants was answers to practice lessons; however the measure of performance was the total score on the NDRT. Administering several versions of the NDRT periodically throughout the experiment so that feedback would have been consistent with the goal might have been more appropriate.

When individuals commit to a goal, perceived negative discrepancies between goals and performance may serve as an additional motivator. However, at some point the negative discrepancy can become of sufficient magnitude to reduce performance motivation (Bandura & Cerone, 1983). The negative discrepancy (i.e., the scores on the practice exams) was not measured in this study, and could be a contributing factor to the non-significant linkage between feedback and implementation effort. Also the self evaluation manner in which the feedback was received may have resulted in a different motivational inducement than feedback provided by others. The self evaluation may have been perceived as a learning experience rather than performance feedback and thus the lack of a feedback - effort linkage.

Decision confidence was not affected by how the subject received his/her reading plan. Self-generated computer output was expected to affect decision confidence more than the chauffeur-generated computer output, but the data did not support this contention. Perhaps the level of interaction in the self-generated treatment was not sufficiently high to engage the participants. Another reason for this result may be that these participants saw no difference between generating the computer output themselves or having another (the "chauffeur") generate it for them.

Decision confidence is a critical component of effective decision making and for critical decisions, confidence is no less important than quality in outcome success. Despite this, there has been no integrated study of the relationships between decision confidence, implementation effort, and outcome success. The present study provides empirical evidence that without confidence in the decision, the best laid plans of mice and men may go awry. Confidence influences implementation effort, which in turn affects outcome success. Our results highlight the importance of decision confidence to implementation effort and outcome success. Prior findings that self-efficacy influences goal-setting behavior and goal commitment were confirmed. We additionally demonstrated that goal commitment had a positive impact on implementation effort.

These results have important implications for researchers and practitioners. The importance of decision confidence is apparent as is the need to further investigate its role in implementation efforts and decision outcomes. Individuals in decision-making and decision-implementing positions should be aware of the influence of decision confidence on implementation effort and successful outcomes. This study also elucidates the importance of future study to understanding its role in decision making and outcome success.

REFERENCES

Anderson, J., & Gerbing, D. W. (1988). Structural modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103* (3): 411-423.

Bagozzi, R. P., & Yi, Y. (1989). On the use of structural equation models in experimental designs. *Journal of Marketing Research, 26*: 271-284.

Barnard, C. (1938). *The functions of the executive.* Cambridge, MA: Harvard University Press.

Bandura, A. Self-efficacy: (1977). Toward a unifying theory of behavioral change. *Psychological Review, 84*: 191-215.

Bandura, A., & Cervone, D.(1983). Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems. *Journal of Personality and Social Psychology, 45*:1017-1028.

Beach, L. R. (1985). Action: Decision implementation strategies and tactics. In Frese, M., & Sabini, J. (Eds.), *Goal directed behavior: The concept of action in psychology*, (pp. 123-131). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Beyth-Marom, R., Dekel, S., Gombo, R., & Shaked, M. (1985). *An elementary approach to thinking under uncertainty.* Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Boal, K., & Bryson, J. (1987). Representation, testing and policy implications of planning processes. *Strategic Management Journal, 8*: 211-231.

Bonoma, T., & Clark, B. (1988). *Marketing performance assessment.* Boston, MA: Harvard Business School Press.

Brown, J. I., Bennett, J. M., & Hanna, G. (1981). *The Nelson-Denny reading test: Examiner's manual form E and F.* Chicago, IL: Riverside Publishing Company.

Browne, M., & Cudeck, R. (1993). Alternative ways of assessing model fit. In Bollen, K. A., and Long, J. S. (Ed.), *Testing structural equation models*, Los Angeles, CA:Sage Publications.

Conlon, E., & Wolf, G. (1980). The moderating effect of strategy, visibility, and involvement on allocation behavior: An Extension of Staw's Escalation Paradigm.*Organizational Behavior and Human Decision Performance, 26*:176-192.

Cravens, D. W. (1991). *Strategic marketing.* Homewood, IL: Richard D. Irwin. Dean, J. Jr., & Sharfman, M.. (1996). Does decision process matter? A study of strategic decision-making effectiveness. *Academy of Management Journal, 39*: 368-396.

Digman, L. A. (1990). *Strategic management: Concepts, decision, cases.* Boston, MA: Irwin.

Earley, P. C. (1988). Computer-generated performance feedback in the magazine-subscription industry. *Organizational Behavior and Human Decision Processes, 41*:50-64.

Earley, P. C., Northcraft, G. B., Lee, C., & Lituchy, T. R. (1990). Impact of Process and Outcome Feedback on the Relation of Goal Setting to Task Performance. *Academy of Management Journal*, *33*: 87-97.

Earley, P., Wojnaroski, P., & Prest, W. (1987). Task planning and energy expended: Exploration of How goals influence performance. *Journal of Applied Psychology*, *72*: 107-114.

Einhorn, H. (1986). Accepting error to make less error. *Journal of Personality Assessment*, *50*: 387-395.

Einhorn, H., & Hogarth, R.(1978). Confidence in judgment: Persistence of the illusion of validity. *Psychological Review*, *85*(5), 395-416.

Fornell, C., & Larcker, D.(1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 1981, 39-50.

Fox, F., & Staw, B. (1979). The trapped administrator: The effects of job insecurity and policy resistance upon commitment to a course of action. *Administrative Science Quarterly*, *24*, 449-471.

Frayne, C. & Latham, G.(1987). Application of social learning theory to employ self management of attendance. *Journal of Applied Psychology*, *72*, 387-392.

Freeman, R. (1984). A slip of the chip on computer spreadsheets can cost millions. *The Wall Street Journal*, *205* (35),14.

Guth, W., & MacMillan, I.(1986). Strategy implementation versus middle management self-interest. *Strategic Management Journal*, *7*, 313-327.

Hogarth, R. (1980). *Judgment and choice: The psychology of choice*. New York: John Wiley & Sons.

Hollenbeck, J., Williams, C., & Klein, H. (1989). An empirical examination of the antecedents of commitment to difficult Goals. *Journal of Applied Psychology*, *74*, 18-23.

Huff, S. (1986). DSS development: Promise and practice. *Decision Support Systems*, *1*, 8-21.

Jöreskog, K., & Sörbom, D. (1993) *LISREL 8*. Chicago, IL: Scientific Software Inc.

Katzel, R., & Thompson, D. (1990). Work motivation: Theory and practice. *American Psychologist*, *45*(2), 144-153.

Kasper, G. (1996). A theory of decision support systems design for user calibration. *Information Systems Research*, *7*, 215.

Koontz, H. (1958). A preliminary statement of principles of planning and control. *Journal of the Academy of Management*, *1*, 45-61.

Locke, E., Frederick, E., Lee, C., & Bobko, P.(1984). Effect of self-efficacy, goals, and task strategies on task performance. *Journal of Applied Psychology*, *69*, 241-251.

Locke, E., & Latham, P. (1990). *A theory of goal setting and task performance*. Englewood, NJ: Prentice-Hall, Inc.

Locke, E., Shaw, K., Saari, L., & Latham, G. (1981). Goal setting and task performance:1969-1980, *Psychological Bulletin*, *90*(1), 125-152.

McIntyre, S. (1982). Experimental study of the impact of judgment-based marketing models, *Management Sciences*, *28*(1), 17-33.

Milliken, F., and Vollrath, D. (1991). Strategic decision-making tasks and group effectiveness: Insights from theory and research on small group performance. *Human Relations*, *44*(12), 1229-1253.

Muir, B. (1987). Trust between humans and machines, and the design of decision aids. *International Journal of Man-Machine Studies*, *27*(5), 527-539.

Neale, M., & Bazerman, M.(1985). The effects of framing and negotiator overconfidence on bargaining behaviors and outcomes. *Academy of Management Journal*, *28* (1), 34-49.

Nelson-Denny Reading Test. Chicago, IL: The Riverside Publishing Company, 1981.

Northcraft, G., & Earley, P. (1989). Technology, credibility, and feedback use. *Organizational Behavior and Human Decision Processes*, *44*(1), 83-96.

Rousseau, D. (1995). *Psychological contracts in organizations*. New York: Sage Publications

Russo, J., & Schoemaker, P. (1992). Managing overconfidence. *Sloan Management Review*, *33*(2), 7-17.

Sniezek, J. (1992). Groups under uncertainty: An examination of confidence in group decision making. *Organizational Behavior and Human Decision Processes*, *52*, 124-155.

Sniezek, J. & Henry, R. (1989). Accuracy and confidence in group judgment. *Organizational Behavior and Human Decision Processes*, *43*(1), 1-28.

Staw, B. (1976). Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action. *Organizational Behavior and Human Decision Processes*, *16*(1), 27-44.

Staw, B. (1981). The escalation of commitment to a course of action. *Academy of Management Review*, *6*(4), 577-587.

Stevick, R., Martin, K., & Showalter, L. (1991). Importance of decision and postdecision dissonance: A return to the racetrack. *Psychological Reports*, *69*(2), 420-422.

Thompson, A.A., & Strickland, A. J. (1999). *Strategic management*. New York: Irwin-McGraw Hill.

Trull, S. (1966). Some factors involved in determining total outcome success. *Management Science*, *12*(6), 270-280.

Weick, K. E.(1990). Introduction: Cartographic myths in organizations. In A. S. Huff (Ed.), *Mapping strategic thought* (pp. 1-10). Chichester, UK: John Wiley and Sons.

Wernham, R. (1987). Implementation: The things that matter. In King, W. R., and Cleland, D. I. (Eds.), *Strategic planning and management handbook*. New York: Van Nostrand Reinhold, 439- 477.

Disposition Toward Thinking Critically: A Comparison of Pre-service Teachers and Other University Students

Martin E. Eigenberger
University of Wisconsin-Parkside

Karen A. Sealander
Northern Arizona University

James A. Jacobs
Indiana State University

Suzanne M. Shellady
Central Michigan University

A preliminary study was conducted examining attitudes toward some of the requisite dispositions involved in critical thinking. The Attitude Toward Critical Thinking Scale (CTS) was used to assess dispositions related to critical thinking in a sample of university students. Elementary, secondary, and special education majors, as well as majors classified as Liberal Arts & Sciences, Traditional Social Sciences, and Applied Social and Health Sciences were compared. Results suggested that students sampled from Arts & Sciences and Traditional Social Science are more favorably predisposed to critical thinking principles than are Education and Applied Social and Health Science students. Results are addressed within the context of pedagogical needs of the information age.

Critical thinking skills have received much attention over the past two decades. Most authors who address this subject agree that the ability to think critically and autonomously is the cornerstone of democracy and the antidote to authoritarianism and closed-mindedness (Ennis, 1992; McPeck, 1990; Paul, 1984; Siegel, 1988). It is further suggested that accelerated discovery and information flow necessitate the development of thinking skills that will be adaptive to the new information environment. The method of critical thinking, which finds its typical expression in philosophical analysis, requires a non-didactic, forum-style classroom environment conducive to dialog and Socratic questioning. This kind of teaching and learning is aimed at understanding the

---

Author info: Correspondence should be sent to: Dr. Martin Eigenberger, College of Arts & Sciences, U. of Wisconsin-Parkside, 600 Wood Rd., Kenosha, WI 53141

North American Journal of Psychology, 2001, Vol. 3, No. 1, 109-118.

© NAJP

underlying assumptions and logical structure of subject areas as opposed to a strict delivery of factual content (Paul, 1993). By instilling the methods and habits of critical inquiry, students will be less dependent on 'the facts'- which are constantly changing. Students so educated are more likely to become autonomous thinkers, able to extract reliable information from a wide variety of sources, and able to make more incisive inferences. However, philosophical ways of thinking, much like a second language, are best practiced early and often. Thus critical thinking should be infused across the curriculum beginning in the early elementary classroom (Lippman, 1991). The attitude of future teachers regarding their readiness to implement critical thinking may be regarded as an important and integral consideration of teacher education programs (Presseisen, 1986).

Critical thinking is defined both as a psychological attitude and an intellectual skill. It is an attitude of reflective skepticism and a willingness and ability to suspend judgement pending a critical examination of evidence (McPeck, 1981). Critical thinking is a refusal to accept information without good reasons for belief (Siegel, 1988), and this attitude not only attempts to examine information in analytic bits but also pulls within its scope worldviews and personal perspectives (Paul, 1984). Critical thinking is also a skill that uses informal logic, philosophy, and subject-specific facts to sort out what are adequate, acceptable, or good reasons for belief.

Both the attitude and skill of critical thought are as important today as in the past. Historically, critical thinkers such as Descartes systematized an approach to knowledge acquisition and revolutionized philosophy and empirical science. Scientific methodologies initiated during the Enlightenment period have produced a seeming over-abundance of information. According to Paul (1993) routinized learning, stable world views, and formulaic teaching are fast becoming handicapping conditions for individuals and societies. Accelerating change, Paul suggests, is upon us, as half the content of subject areas becomes obsolete or invalid every six years. We are increasingly being asked to sort through a maze of new or reorganized information. Toffler (1970) notes that from 1776 to 1960 the stores of knowledge doubled. From 1960 to 1965 knowledge again doubled. It is predicted that as we enter the year 2002, factual knowledge will double every 73 days (Toffler, 1990). Additional examples of the enormous amounts of information available include the 1000 new books published every day and the fact that there is more information in a single edition of the New York Times than there was in the lifetime of an individual living in 1776 (Toffler, 1990).

The use of critical thinking skills, or content-independent reason, is the only practical way to sort out which information may be valid. As in the past, critical thinking is necessary to guide social changes as well as those changes proposed and presented by today's educational reform movement. If critical thought is not employed, then the field of education may be more likely to be driven by political agendas and market forces rather than real cognitive and informational needs that are relavent to the adaptability of individuals and institutions. Hiring practices of corporations like Motorola and Honeywell are beginning to address the need for critical thinkers. Companies like these emphasize the requirement for potential employees to be able to independently process and use new information as well as work with others to creatively use the information.

Given the information explosion, the notion that knowledge is power, and the needs of employers regarding the individuals they hire, critical thinking is fast becoming a necessary survival skill. Although the acquisition of adaptive skills may be related to intelligence, the learning of critical thinking as an adaptive behavior requires practice on the part of the learner. Even if there are "naturals" or native critical thinkers as there are natural artists, guided practice of critical thinking skills is still essential. It is logical then, that critical thinking be encouraged in the classroom. This encouragement may take the form of emphasizing the need for good reasons rather than just correct answers, and insisting on evidence rather than opinion. The questioning of motives and assumptions, and the exploration of implications must be practiced. The practice of critical thinking also requires the acceptance and respect of divergent perspectives (Marzano, 1992; Slavin, 1994). Attention therefore, to the nature and intellectual ecology of the classroom is essential, as it might well enhance or impede the development of critical thinking.

To adequately deal with critical questions, the learning outcomes sought by both the teacher and the student must be extended beyond merely finding test-relevant answers, to one of exploration and reflection. To achieve these outcomes it is essential for both the teacher and students to tolerate the ambiguity and periodic confusion that often accompanies the process of discovery via critical thinking. The classroom would seem to be the ideal arena in which to nurture, practice, and implement thinking skills (Hester, 1994).

Student use of critical thinking in the classroom, although pedagogically sound and seemingly widely accepted, can be a double-edged sword. On the one side, the students are using skills which more fully engage the mind in the assimilation and understanding of subjects and subject matter. On the other side, there is the attitude of skepticism and questioning, which may have the appearance of challenging authority, often including the teacher's authority (Bornstein, et al, 1977). Hence, for some teachers critical thinking may be controversial. The teacher with an authoritarian turn of mind, who works for conformity and regularity in his or her classroom, may have something less than a full appreciation for students who are asking critical questions.

Indeed, that teacher may view the student as threatening, non-compliant, and rude (Drabman & Patterson, 1981). It would seem that the teacher's assumptions about the nature of authority and skepticism, and the value of justified beliefs, may affect the quantity and quality of critical thinking that actually takes place in the classroom.

The central question of this study was: Do future educators value the critical thinking process and the behavioral attributes of the critical thinker, or are they more likely to ignore, limit or extinguish such behaviors? In an attempt to begin addressing this question, students enrolled in institutions of higher education volunteered to complete the Attitude Toward Critical Thinking Scale (Eigenberger, Sealander, & Seckinger, 1996). The purpose of this preliminary investigation was to assess the degree to which critical thinking attitudes are present in future educators as compared to other university students.

METHOD

Participants and Procedure

Participants in this study were 486 students enrolled in undergraduate and graduate courses in the College of Education and three other areas of academic discipline. Other areas were classified as, a) Traditional Liberal Arts and Sciences (A&S), consisting mainly of history, philosophy, English, fine arts, and natural science majors, b) Traditional Social Science (TSS), which was primarily made up of psychology, sociology, and political science majors, and c) Applied Social and Health Sciences (ASHS), composed of majors from fields such as criminal justice, social work, nursing, speech pathology, occupational therapy, and nutrition science. The survey procedure netted participants from other disciplinary areas such as Engineering and Business, but total numbers in these categories were trivial and were dropped from the analysis. A total of 22 participants declared themselves as Undecided, and were included as a group in the analysis.

Participants were volunteers drawn from courses at three universities in the Western and Mid-western states. The largest sample of participants was drawn from Colleges of Education and consisted of 240 students. All of the participating institutions offered four-year, undergraduate degrees, as well as post-graduate degrees in Education. Of the Education students, 114 participants (96 women, 18 men) were elementary education majors, 69 (32 women, 37 men) were secondary education majors, and 58 (48 women, 10 men) were special education majors. These categories within Education were examined to test the hypothesis that the critical thinking disposition of secondary education majors would differ significantly from other majors within Education. The total Education sample contained 176 women and 64 men. Mean age of the participants was 23.4 years. Ninety-five participants were A&S majors, consisting of 36 women and

59 men. Traditional Social Science consisted of 83 participants, of which 45 were women and 38 were men. Applied Social & Health Sciences was made up of 46 participants - 28 women and 18 men. Of the students in the Undecided category, there were 13 women and 9 men.

The CTS was administered along with a demographics page, by several professors who were teaching courses in a variety of disciplines. Participants were offered extra credit for completing the scale and were given class time for this purpose.

Instrumentation

Items for the CTS were developed by the authors, and were derived deductively from conditions assumed necessary for effectively engaging in critical and reflective thought (e.g., Dewey, 1933; Paul, 1993; Russell, 1912). These necessary conditions were operationalized as statements or 'sentiments' reflecting traits such as intellectual openness, faith in reason, curiosity, respect for science, willingness to suspend judgement, and independent thinking. The CTS was designed to sample an attitude spectrum incorporating 'pro' and 'con' dispositions toward critical thinking. It was developed as an alternative to commercially available scales such as the California Critical Thinking Disposition Inventory (Facione & Facione, 1992), which is a comprehensive, reliable, but rather lengthy, multidimensional instrument. The CTS is a relatively short scale, intended for use in omnibus survey packets, or, for example, as an expedient indicator of the strength of the critical thinking disposition in an individual classroom. The CTS contains 15 items scaled on a 7-point disagree-agree Likert-type format. The scale is "balanced," with 7 items worded in the *pro* direction, capturing a positive attitude toward the principles of critical thinking, and 8 items worded in the *con* direction, indexing negative attitudes toward the requisite conditions of critical thinking. Items typical of the CTS are "One good way to find what is true, is to doubt everything until there are good logical or scientific reasons to believe it"; "Cooperation and respect for tradition are more important virtues than being a critical thinker"; and, "In the quest for what is really true, one should be prepared to doubt and perhaps discard any current beliefs about oneself, God, and the world." The CTS has a possible raw score range from 15 to 105 points. The fifteen items comprising the CTS are listed in the Appendix.

Evidence of construct validity for the CTS was collected in a separate pilot study (Eigenberger, Sealander, & Seckinger , 1996) and used three measures. The CTS correlated with the following: Troldahl and Powell's (1972) version of the Dogmatism Scale (r = -.54), Altemeyer's (1988) Right-Wing Authoritarianism Scale (r = -.58), and summed scores from the Truth-seeking, Open-mindedness, and Inquisitiveness sub-scales of Facione & Facione's (1992) California Critical Thinking Disposition

Inventory ($r = .68$). A Cronbach's alpha reliability coefficient of .83 was established for the CTS in both the pilot and the present study.

Factor Analysis

A principal components factor analysis was conducted followed by varimax rotation. An initial solution produced 3 factors, the first of which had an eigenvalue of 4.78 and accounted for 31.9% of the variance. The second factor had an eigenvalue of 1.67 and accounted for 10.6% of the variance. A third factor had an eigenvalue of 1.12 and accounted for 7.5% of the variance. Examination of the scree plot however, indicated the presence of two interpretable factors. A two-factor solution in fact provided the most parsimonious account of scale structure. Items loading onto Factor 1 seemed to relate to teaching or promoting critical thinking versus respect for tradition and practicality. Items loading on Factor 2 related to a more general skepticism versus varied categories of belief. Abbreviated scale items and factor loadings are displayed in Appendix B.

RESULTS

Two one-way analyses of variance (ANOVA) were conducted. The ANOVAs examined CTS scores between the academic areas, Education, A&S, TSS, ASHS, along with students classified as Undecided. The analysis of academic areas suggested that the groups differed significantly on CTS scores [$F(4, 480) = 17.9$, $p < .001$]. Post hoc comparisons using the Tukey HSD procedure further indicated that A&S students and TSS students scored significantly higher than Education majors, ASHS majors, and Undecided students.

TABLE 1 Summaries of CTS Scores Among Academic Areas

	M	SD	N
Education	64.5	13.4	240
Elementary	62.7	11.4	114
Secondary	67.9	15.3	69
Special Education	63.6	14.1	57
A&S	77.0	15.3	95
TSS	72.9	13.0	83
ASHS	64.0	13.3	46
Undecided	65.1	12.5	22
Total	68.4	14.6	486

A second one-way ANOVA compared scores of three sub-groups within the College of Education and the other four non-education academic groups. Within the College of Education, three levels of specialization were selected: 1) elementary, 2) secondary, and 3) special education. Results suggested significant differences in mean CTS scores between the groups. [F (6, 479) = 13.3, $p < .001$]. A follow-up Tukey HSD test revealed secondary education majors had significantly higher CTS scores than elementary education majors, but did not differ significantly from either special education majors, ASHS majors, or Undecided students. Secondary education students scored significantly lower than both A&S and TSS majors. Means and standard deviations of all groups are presented in Table 1.

A small but statistically significant difference between males and females was observed in mean CTS scores as measured by a t-test. The male respondents (M = 71.3, SD = 16.4) scored higher than female respondents (M = 66.5, SD = 13.0), t (482) = -3.55, $p < .01$.

DISCUSSION

The results of this study, are best viewed as exploratory, with limitations on generalizability. Nevertheless, within these limitations, the results indicated significant differences existed in critical thinking dispositions among several identifiable academic groups within a university population. The most salient observation was the significant differences in CTS scores between Education students and both traditional liberal arts students and traditional social science students. Regarding critical thinking dispositions, students enrolled in Education, with the exception of secondary education majors, more closely resembled students preparing to enter occupations such as law enforcement, probation and parole, social work, physical therapy and so on. These occupations, while necessary and valuable, are not generally identified with the promotion of critical thinking-related dispositions such as intellectual openness, wide-ranging curiosity, autonomy, and willingness to suspend judgement.

This study also suggested a difference in critical thinking attitude between majors within the College of Education. It was observed that secondary education majors scored significantly higher on the CTS than their elementary education counterparts. They also scored higher, albeit not significantly higher, than special education and ASHS students. This observed differential within Education may be due in part to the secondary education major's tendency to be content specialists, perhaps with a deeper relationship to their subject areas. Secondary education majors may also enter their specialties because they have a preference for working with students who are generally functioning at more abstract cognitive levels, where the ability to perform conceptual analysis is possible.

While a gender differential relative to CTS scores was observed, it should be noted that this difference may be an artifact of the sample, and explained by the large ratio of female to male students sampled in the College of Education – students who tend to have lower CTS scores. At any rate, the disproportionate representation of female students in the sample should not produce a great deal of confidence in the stability of the *t*-ratio.

Given the exploratory nature of the study, alternative explanations for the observed results could also be considered. For instance, the subjects sampled from the College of Arts & Sciences may not have been representative. As noted, majors from traditional humanities courses were sampled, i.e., philosophy and history. Students majoring in these areas can be *expected* to have a strong critical thinking orientation, since their course work and assessment procedures center on analytic abilities and wide general interests. The effect of sampling these majors in high proportions may be to bias CTS means for the College of Arts & Sciences as a whole. However, this concern is muted by similar CTS scores of students who were classified as Traditional Social Science majors.

If the observed differences in critical thinking disposition scores between Education students and other academic areas is stable, then obvious questions in this regard arise: "Is this difference between majors a source of concern?" If some Education majors value critical thinking dispositions less than Arts & Sciences or Traditional Social Science students, then are primary and secondary American schools employing teachers who are not as prepared as they should be regarding those intellectual dispositions necessary for critical thinking? What effect might this intellectual disposition have in the way children are taught in the classrooms? Will those teachers who are not predisposed to value and therefore implement critical thinking, regard the questioning and skeptical student as problematic? Would teachers so disposed, view the exploratory critical process as a distraction from content coverage? On the other hand, if teachers value the critical thinking attitude, will they be more tolerant of student questions and open to using the logic of subject areas as an instructional strategy (e.g. Paul, 1993)? Would the introduction of critical thinking attitudes and behaviors such as reflective skepticism and Socratic questioning into a classroom result in a more productive and ecologically stable classroom?

Given the potential importance of pre-service teacher attitudes on educating students for engaging in critical discourse and democratic living, it is evident that such attitudes should be further examined relative to their possible interaction with reform-oriented pedagogical changes. Should further studies confirm future educators' relative disinterest in, or discomfort with critical thinking-related dispositions, then more attention to 'thinking skills' acquisition may be needed in certain of our teacher preparation programs.

REFERENCES

Altemeyer, B. (1988). *Enemies of freedom.* San Francisco: Josey-Bass.

Bornstein, M.R., Bellack, A.S., & Hersen, M. (1977). Social skills training for unassertive children: A multiple baseline analysis. *Journal of Applied Behavior Analysis, 10,* 183-195.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 30,* 1-13.

Dewey, J. (1933). *How we think.* New York, Houghton Mifflin.

Drabman, R.S., & Patterson, N.J. (1981). Disruptive behavior and the social standing of exceptional children. *Exceptional Children Quarterly, 1* (4), 44-55.

Eigenberger, M.E., Sealander, K.A., & Seckinger, D., (1996). Validation study of three student perception inventories. *Univer. of Wyoming College of Education and Wyoming School Partners Final Report.*

Ennis, R.S. (1995). *Critical thinking.* Upper Saddle River, NJ: Prentice-Hall.

Facione, P.A., & Facione, N.C. (1992). *The California critical thinking disposition inventory.* Millbrae, CA: The California Academic Press.

Hester, J. (1994). *Teaching for thinking: A program for school improvement through teaching critical thinking across the curriculum.* Durham, NC: Academic Press

Lippman, M. (1991). *Thinking in education.* New York: Cambridge Univer. Press.

Marzano, R.J. (1992). *A different kind of classroom: Teaching with dimensions of learning.* Alexandria, VA: Association for Supervision and Curriculum Development.

McPeck, J. (1981). *Critical thinking and education.* New York: St. Martin's Press.

McPeck, J. (1990). *Teaching critical thinking.* London, Routledge.

Paul, R.W. (1984). Critical thinking: Fundamental to education in a free society. *Educational Leadership, 42,* 1-14.

Paul, R.W. (1993). The logic of creative and critical thinking. *American Behavioral Scientist, 37,* 21-39.

Paul, R.W. (1995). *Critical thinking.* Santa Rosa, CA: Foundation for Critical Thinking.

Presseisen, B. (1986). *Thinking skills: Research and practice.* Washington, DC: National Education Association.

Russell, B. (1912). *The problems of philosophy.* London: Oxford Univer. Press.

Schoenfeld, A.H. (1982). Measures of problem-solving instruction. *Journal for Research in Mathematics Education, 13,* 31-49.

Seigel, H. (1988). *Educating reason.* London: Routledge.

Slavin, R.E. (1994). *Educational psychology: Theory and practice* (4th ed.). Needham Heights, MA: Allyn & Bacon.

Toffler, A. (1970). *Future shock.* New York: Random House.

Toffler, A. (1990). *Powershift: Knowledge, wealth and violence at the edge of the 21st Century.* New York, Bantam.

Troldahl, V.C., & Powell, F.A. (1972). A short-form dogmatism scale for use in field studies. *Social Forces, 44*, 211-214.
Wasserman, S. (1987). Teaching for thinking: Louis E. Raths revisited. *Phi Delta Kappan, 68* (6), 460-466.

APPENDIX A

Sample Items Comprising the Attitude Toward Critical Thinking Scale

2. Teaching students good "solid basics" like the 3R's is more important than having them question and critique everything they are told.
3. *All* our beliefs are subject to criticism and change - human beings have no access to a "perfect truth."
5. A child should not be taught "critical thinking" if it has the potential to upset traditional values.
6. If old beliefs about the world and mankind can't stand up to modern science and criticism, then we ought to abandon such beliefs.
8. There should be *no* limits on what is doubted and criticized, even if those things are our society's most deeply held or "sacred" beliefs.
10. *All* forms of political and religious authority must be constantly questioned and challenged or else we will have more oppression and injustice in society.
12. Belief in one's God and country should not be criticized or totally doubted.
13. Critical thinking is fine but it doesn't really prepare students for life in the real world.

APPENDIX B

Factors and Loadings of the Attitude Toward Critical Thinking Scale Items

Factor	Loading
Factor 1	
There is a limit to criticizing normal ways of doing things.	.76
A child should not be taught critical thinking if it upsets tradition.	.75
Cooperation and respect more important than critical thinking.	.69
Critical thinking does not prepare students for real life.	.63
Teaching students "solid basics" more important than questioning.	.64
Critical thinking in the classroom takes a lot of time and energy.	.57
We should just accept laws and moral teaching without skepticism.	.48
Belief in one's God and country should not be criticized.	.43
Factor 2	
One should be prepared to discard beliefs.	.69
Political and religious authority must be questioned.	.68
Abandon old beliefs about that can't stand up to criticism.	.67
Doubt everything until there are good reasons to believe it.	.54
It is better to be a true believer than a critical thinker.	.50
There should be no limit on what is doubted and criticized.	.48
All out beliefs are subject to criticism and change.	.36

Age-discrepant Relationships: Do These Romances Fare Well?

Ann Zak, Erica Armer, Kristina Edmunds
Tara Fleury, Maria Sarris, Beth Shatynski
College of Saint Rose

Though society and relationship researchers alike often view age-discrepant romances as "taboo," we expected, in line with divorce rates, that age-discrepant relationships would fare well. We hypothesized that persons in age-dissimilar relationships would score higher on measures of trust and unselfish love and lower on measures of jealousy than those in age-similar relationships. Results confirmed predictions.

One clear indicator of healthy relationship functioning is similarity (Byrne, 1971). Persons of similar age, race and attitudes are more likely to become attracted to one another and eventually marry than dissimilar others (Doosje, Rojahn and Fischer, 1999; Waris, 1997). Indeed, age-discrepant relationships are often considered "taboo" and doomed to fail (Shehan, Berardo, Vera and Carley, 1991). With public reaction to May-December romances (such as those uniting Catherine Zeta-Jones with Michael Douglas and Anna Nicole Smith with J. Howard Marshall) largely negative, it is almost surprising that age-dissimilar partners attract. Once the relationship is established, however, differences in age may enhance relationship quality.

Gutfield (1995), for example, cited national divorce rates that drop significantly among age-dissimilar spouses. He claimed that age-discrepant intimates are better at anticipating difficulties and resolving issues regarding love, sex, family and finances. Consistent with sociobiological theory, it was also expected that men would enter into relationships with younger, more fertile women and women would seek older, more financially secure men to ensure the survival of their offspring (Hayes, 1994). Thus, we believe that age-dissimilar relationships may be successful. In particular, we predicted that persons in age-dissimilar relationships would report more trust and unselfish love and less jealousy than those in age-similar relationships.

Author info: Correspondence should be sent to: Dr. Ann Zak, Dept. of Psychology, College of St. Rose, 432 Western Ave., Albany, NY 12203. Zaka@strose.edu

North American Journal of Psychology, 2001, Vol. 3, No. 1, 119-122.

© NAJP

METHOD

324 participants were recruited at area colleges and public places in the Northeastern U.S. Age of participants ranged from 18 to 71 years (M = 24). 249 participants were seriously dating, 11 were engaged and 64 were married. Reported age differences between participants and their current partners ranged from 0 to 20 years (M = 3). Intimates completed the Trust Scale (Rempel, Holmes & Zanna, 1985), the Agape (unselfish love) subscale of the Love Attitudes Scale (Hendrick & Hendrick, 1986) and the Jealousy Scale (Strzysewski and Comstock, 1991) and were then debriefed.

RESULTS AND INTERPRETATION

A median-split was performed on age differences between participants and current partners, with 0-3 years classified as age-similar and 4 or more years as age-dissimilar. The median age discrepancy in the age-dissimilar group was 5.4 years, as compared to 1.3 years in the age-similar group. One-way analyses of variance were then conducted with total trust, agape and jealousy scores as dependent variables and age differences as the independent variable. As expected, partners in age-dissimilar relationships reported more trust, F (1, 322) = 4.3, $p<$ .05, more agape (unselfish love), F (1,322) = 9.2, $p<$ .001, and less jealousy, F (1, 322) = 11.4, than partners in age similar relationships (see Table 1).

TABLE 1 Total Trust, Unselfish Love and Jealousy Scores (SD) by Age-Differences

	Trust	Agape*	Jealousy
Age-Dissimilar Relationships	95.3	13.5	23.8
	(*SD*=4.5)	(*SD*=2.3)	(*SD* = 2.9)
Age-Similar Relationships	90.4	14.7	27.4
	(SD=3.7)	(SD=2.9)	(*SD* = 3.1)

*Lower Agape totals indicate higher reported unselfish love

Overall, the findings indicate that partners in age-dissimilar relationships trust their partners more, report more unselfish, sacrificing love and experience less jealousy than those who find romance within their cohort.

A partial explanation for the results stems from the finding that couples in the age-dissimilar group were about four years older than couples in the age-similar group. If age is liked to maturity, and maturity is linked to relationship quality, then age-dissimilar relationships might work partly because at least one member of each age-discrepant couple is likely to be older than the average age for couples involved in serious relationships.

In the future researchers may find it fruitful to examine whether persons in such "taboo" relationships have anticipated relationship conflict and resolved key issues. In other words, what are some other reasons why age-dissimilar relationships work?

REFERENCES

Aune, K. S., & Comstock, J. (1991). Experience and expression of jealousy comparisons between friends and romantics. *Psychological Reports, 69*, 315-319.

Byrne, D. (1971). The attraction paradigm. New York: Academic Press.

Doosje, B., Rojahn, K., & Fischer, A. (1999). Partner preferences as a function of gender, age, political orientation, and level of education. *Sex Roles, 40*, 45-60.

Gutfield, G. (1995). Older men, younger women. *Men's Health, 10*, 39-41.

Hayes, A. F. (1994). Age preferences for same and opposite sex partners. *Journal of Social Psychology, 135*, 125-133.

Hendrick, C., & Hendrick, S. S. (1986). A theory and method of love. *Journal of Personality and Personal Relationships, 5*, 161-183.

Rempel, J. K., Holmes, J. G., & Zanna, M. (1985). Trust in close relationships. *Journal of Personality and Social Psychology, 49*, 95-112.

Shehan, C. L., Berardo, F. M., Vera, H., & Carley, S. M. (1991). Women in age-discrepant marriages. *Journal of Family Issues, 12*, 291-305.

Strzysewski, A., & Comstock, J. (1991). Experience and expression of jealousy: Comparisons between friends and romantics. *Psychological Reports, 69*, 315-319.

Waris, R. G. (1997). Age and occupation in selection of human mates. *Psychological Reports, 80*, 1123-1126.

The Cognitive Style Index and the Legal Profession in Nova Scotia: A Factor-analytic Study

H. Joseph Murphy
Dalhousie University

Pauline A. Doucette
William E. Kelleher
John G. Reid
Jeffrey D. Young
Mount Saint Vincent University

A factor analysis of scores on the Cognitive Style Index (CSI) for a sample of 615 lawyers in eastern Canada was completed. The findings of the present study are consistent with the hypothesis that there is a single dimension underlying cognitive style as measured by the CSI.

Cognitive styles generally refer to the stable and habitual manner in which people process information. Messick (1976) suggested that individuals have preferences for organizing information and experience and, consistent individual differences in these ways of organizing and processing information and experience have come to be called cognitive styles. They are conceptualized as stable attitudes, preferences, or habitual strategies determining a person's typical modes of perceiving, remembering, thinking, and problem solving (p. 5).

Further, Messick (1976) suggested that cognitive styles are consistent in manner, develop slowly and experientially and are not readily modified by training. They are stable and pervasive across behavioral situations and form a part of an individual's personality. They differ from ability in that ability focuses on the "what" of cognition whereas cognitive styles focus on the "how." In general, cognitive styles are considered bi-polar, with individuals falling somewhere between the extremes. Each of the extremes has different implications for cognition and each extreme has value depending on the situation.

The cognitive styles of various pre-professional groups have been explored, for example, information management students (Casey,

---

Author info: Correspondence should be sent to: Dr. Jeffrey D. Young, Mt. Saint Vincent U., Halifax, Nova Scotia B3M 2J6.

North American Journal of Psychology, 2001, Vol. 3, No. 1, 123-130.
© NAJP

Murphy, & Young, 1996), dental students (Chaytor, Murphy, Boyd, & LaFleche, 1991), and law students (Doucette, Kelleher, Murphy, & Young, 1998b; Townsend & Ede, 1985). Others have examined cognitive style among professional groups, for example, financial analysts (Mykytyn, 1989), organizational executives (Nutt, 1993), and practising lawyers (Doucette, Kelleher, Murphy, Reid, & Young, 1999).

Some of the more widely used measures of cognitive style have been criticized on several grounds. For example, Taylor (1989) and Payne (1987) are critical of Kirton's (1976) measure of adaptation-innovation and Atkinson (1988) has questioned the psychometric soundness of the Learning Style Inventory (Kolb, 1974). The Group Embedded Figures Test (Witkin, et al., 1971) has been criticized because it measures analytical ability not cognitive style (Widiger, Knudson, & Porter, 1980). Also, the utility of the Myers-Briggs Type Indicator (Myers, 1976) for large-scale organizational studies has been questioned because of the large number of items used and the time required to complete it (Allinson & Hayes, 1996).

In an attempt to overcome some of the perceived difficulties with the above measures of cognitive style, Allinson and Hayes (1996) developed the Cognitive Style Index (CSI). The CSI was specifically designed as an easily-administered and easily-scored instrument for use in large-scale organizational studies (Allinson and Hayes, 1996). Similar to the work of Miller (1987) and Nickerson, Perkins, and Smith (1985), Allinson and Hayes (1996) conceptualize cognitive style as a single dimension with an intuitive style at one end and an analytical style at the other end. Intuition "refers to immediate judgement based on feeling and the adoption of a global perspective" and analysis "refers to judgement based on mental reasoning and a focus on detail" (Allinson & Hayes, 1996, p. 122). People tend to favour one cognitive style over another and their style may fall along a continuum between the two end points of intuition and analysis.

A preliminary factor analysis of the Cognitive Style Index was completed to examine the assertion of Allinson and Hayes (1996) that cognitive style is unidimensional. Members of the legal profession were selected because they represent a large and important professional group. Further, this group is quite different in terms of its training and professional practise from the groups of management students and organizational managers used by Allinson and Hayes (1996). This addresses the request of Allinson and Hayes (1996) for "Application of the CSI across a range of organizations and nations ..." (p.131).

METHOD

Participants

A total of 615 members of the Bar Society of Nova Scotia completed the CSI. There were 387 (63%) men who ranged in age from 25 to 85 years with a mean age of 43.8 years (*SD*=9.9). There were 228 (37%) women ranging in age from 25 to 76 years, with a mean of 37.2 years (*SD*=7.8).

Instrument

The CSI consists of 38 items, for example, "I find detailed, methodical work satisfying," to which subjects respond True (T), False (F), or Uncertain (?). Depending on the statement, a score for each statement can be "0," "1," or "2," consequently total individual scores can range from 0 - 76. A low score is indicative of a more intuitive style and a high score is indicative of a more analytical style. Although the instrument is relatively new there is support for its psychometric soundness. For example, Allinson and Hayes (1996) reported a coefficient of stability for test-retest of the CSI of .90, with four weeks between tests, for a group of management students. Doucette et al. (1998a) reported internal consistency coefficients ranging from .84 to .87 for a sample of law school students. Doucette, et al.(1998b) reported internal consistency coefficients ranging from .75 to .88 and a coefficient of stability for test-retest of the CSI of .89, with three weeks between tests, for a sample of business undergraduates.

Allinson & Hayes (1996) suggested that the CSI has concurrent validity because it was capable of discriminating between groups that were presumed to differ in cognitive style. Research on gender differences in cognitive style has been mixed. For example, Agor (1986) suggested that women tended to be more intuitive than men, whereas Kirton (1989) reported that men might be more intuitive than women. The results of Allinson and Hayes (1996) showed that women scored higher on the CSI (i.e., had a more analytical style) than did men, in four out of five samples. Further, women scored significantly higher on the CSI than did men in a study of law students (Doucette, et al., 1998a) and in a study of business administration undergraduates (Doucette et al., 1998b). However, MacGillivary, Murphy, Reid, and Young (1999) found CSI scores for women co-op undergraduates did not differ significantly from those of men co-op undergraduates.

Based on selective recruitment and socialization Allinson & Hayes (1996) suggested that cognitive styles might differ not only between organizations but also between functions within organizations. The significant differences in CSI scores among several functional groups within a broader group of practising managers was offered as evidence of the concurrent validity of the scores produced by the CSI. Further, Townsend and Ede (1985) reported that law students who preferred prosecution as a field of practise were more analytical in style than those who preferred defence as a field of practise. Using the CSI, Doucette, et al. (1998b) examined the cognitive styles of Canadian law students and found that those who preferred the field of litigation scored significantly higher than those who preferred the field of corporate/commercial law. This finding is also supportive of the concurrent validity of CSI scores.

Procedure

The CSI was included with a cover letter and a self-addressed, stamped envelope in a regular monthly information package distributed by the Nova Scotia Bar Society to its 2200 members. Without any follow up 615 responses were received, a response rate of 28%.

RESULTS AND DISCUSSION

Table 1 presents means and standard deviations by item for the Index.

TABLE 1 Means and standard deviations by item for the Cognitive Style Index

Item	M	*SD*	Item	M	*SD*
1	0.88	0.94	20	1.17	0.90
2	1.30	0.91	21	1.33	0.85
3	1.48	0.82	22	1.51	0.75
4	0.99	0.91	23	0.92	0.89
5	1.42	0.80	24	1.23	0.88
6	1.22	0.91	25	1.69	0.55
7	0.79	0.93	26	0.92	0.88
8	1.37	0.83	27	1.22	0.77
9	1.44	0.85	28	1.20	0.87
10	1.46	0.81	29	1.21	0.89
11	1.10	0.92	30	0.62	0.83
12	1.03	0.91	31	1.05	0.93
13	1.55	0.79	32	1.36	0.81
14	1.24	0.88	33	1.83	0.47
15	1.68	0.65	34	0.73	0.86
16	1.56	0.72	35	1.54	0.74
17	0.96	0.87	36	1.69	0.61
18	0.84	0.94	37	0.99	0.90
19	1.11	0.92	38	0.80	0.89

Following the strategy of Allinson and Hayes (1996), parcels of inter-correlating items were created by means of a Principal Components

Analysis performed on the correlation matrix of the 38 items in the CSI (this matrix is not reported here but is available upon request). This is done in order to overcome the tendencies of inter-item correlations to be low (Kline, 1993) and for individual items to be unstable as single variables (Nunnally, 1978). This parcelling process creates objective and homogeneous parcels (Cattell, 1973) which brings "factors by the more exact rotation then possible to greater conceptual precision." (Cattell, 1974, p.103).

The scree plot indicated that approximately five, six or seven-parcel solutions were appropriate, however upon analysis of the respective solutions six was considered a more efficient and interpretable solution. The six parcels account for 39.7% of the total variance in the original 38 items. We note that a Principal Factor Analysis yielded similar interpretations, as expected, since the number of variables exceeded 30 (Gorsuch, 1983).

Table 2 presents the inter-correlations among the six parcels. An exploratory factor analysis was completed on this matrix and the results, without rotation, identified one dominant underlying factor which accounted for 49.1% of the variance in the parcels. Table 3 presents the factor loadings in the one-factor solution and the associated eigenvalue, and the eigenvalues and cumulative explained variance for the factor analysis of the six parcels.

TABLE 2 Intercorrelations Among the Six Parcels

Correlations

	Parcel 1	Parcel 2	Parcel 3	Parcel 4	Parcel 5	Parcel 6
Parcel 1	--					
Parcel 2	42	--				
Parcel 3	47	36	--			
Parcel 4	47	46	49	--		
Parcel 5	33	40	38	41	--	
Parcel 6	32	21	38	39	33	--

Note: Decimal points removed from all coefficients

The findings of the present study are consistent with the hypothesis that there is a single dimension underlying cognitive style as measured by the CSI.

TABLE 3 Loadings on 1st Factor, Associated Eigenvalues and Cumulative Explained Variance for the Factor Analysis of the Six Parcels

Parcel	Loadings on first factor
Parcel 1	72
Parcel 2	68
Parcel 3	74
Parcel 4	78
Parcel 5	67
Parcel 6	60

Eigenvalue 2.94

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Eigenvalue	2.94	0.81	0.70	0.56	0.52	0.47
Cumulative	.49	.63	.74	.84	.92	1.00

REFERENCES

Agor, W. H. (1986). *The logic of intuitive decision making: A research-based approach for top management.* New York: Quorum Books.

Allinson, C.W., & Hayes, J. (1996) The cognitive style index: A measure of intuition-analysis for organizational research. *Journal of Management Studies, 33*, 119-135.

Atkinson, G. (1988). Reliability of the *Learning Style Inventory* - 1985. *Psychological Reports, 62*, 755-758.

Casey, B. Murphy, H. J., & Young, J. D. (1996). Field dependence/independence and undergraduates' academic performance in an information management program. *College Student Journal, 31* (1), 45-50.

Cattell, R. B. (1973). *Personality and mood by questionnaire.* San Francisco: Jossey-Bass.

Cattell, R. B. (1974). Radial parcel factoring-vs-item factoring in defining personality structure in questionnaires: Theory and experimental checks. *Australian Journal of Psychology, 26*(2), 103-119.

Chaytor, D. V., Murphy, H. J., Boyd, M., & LaFleche, R. (1991). *A multi-centre study of student performance on the Group Embedded Figures Test and the Impression/Die Matching Test.* Paper presented at the annual meeting of the American Association of Dental Schools.

Doucette, P. A., Kelleher, W. E., MacGillivary, C. A., Murphy, H. J., Reid, J. G., & Young, J. D. (1999). *Practising lawyers in Nova Scotia: Cognitive style and preferences for practice*. Proceedings, Atlantic Schools of Business Conference, Halifax, Nova Scotia, Canada.

Doucette, P. A., Kelleher, W. E., Murphy, H. J., & Young, J. D. (1998a). Cognitive style and law students in eastern Canada: Preliminary findings. *College Student Journal, 32* (2), 206-214.

Doucette, P. A., Kelleher, W. E., Murphy, H. J., & Young, J. D. (1998b). Test-retest reliability and construct validity of the Cognitive Style Index for business undergraduates. *Psychological Reports, 82*, 595-600.

Gorsuch, R. L. (1983). *Factor analysis*. Hillsdale, NJ: Erlbaum.

Hayes, J., & Allinson, C. W. (1994). Cognitive style and its relevance for management practice. *British Journal of Management, 5*, 53-71.

Kirton, M. J. (1976). Adaptors and innovators: A description and measure. *Journal of Applied Psychology, 61*, 622-629.

Kirton, M. J. (1989). Adaptors and innovators at work. In M. J. Kirton (Ed.), *Adaptors and innovators: Styles of creativity and problem solving* (pp.56-78). London: Routledge.

Kline, P. (1993). *The handbook of psychological testing*. London: Routledge

Kolb, D. A. (1976) *The learning style inventory*: Technical manual. Boston: McBer and Co.

MacGillivary, A. C., Murphy, H. J., Reid, J. G., & Young, J. D. (1999).Cognitive style differences between cooperative and non-cooperative education undergraduates: Some preliminary findings. *The Journal of Cooperative Education, 34*(1), 54-61.

Messick, S. (1976). Personality consistencies in cognition and creativity. In S. Messick & Associates (Eds.), *Individuality in learning* (pp. 4-22). San Francisco: Jossey-Bass.

Miller, A. (1987). Cognitive styles: An integrated model. *Educational Psychology, 7*, 251-268.

Myers, I. B. (1976). *Introduction to type*. Palo Alto: Consulting Psychologists Press, Inc.

Mykytyn, Jr., P. P. (1989). *Group Embedded Figures Test* (GEFT): individual differences, performance, and learning effects. *Educational and Psychological Measurement, 49*, 951-959.

Nickerson, R., Perkins, D., & Smith, E. (1985). *The teaching of thinking*. Hillsdale, NJ: Erlbaum.

Nunnally, J. (1978). *Psychometric theory*. New York: McGraw-Hill.

Nutt, P. (1993). Flexible decision styles and the choices of top executives. *Journal of Management Studies, 30*(5), 695-721.

Payne, R. L. (1987). Individual differences and performance among R&D personnel: Some implications for management development. *R&D Management, 7*, 153-161.

Taylor, W. G. K. (1989). The Kirton Adaptation-Innovation Inventory: A re-examination of the factor structure. *Journal of Organizational Behavior, 10*, 297-307.

Townsend, M. A. R., & Ede, J. (1985). Cognitive style of law students: Prosecution and defense. *Perceptual and Motor Skills, 57*, 762.

Widiger, T., Knudson, R., & Porter, L. (1980). Convergent and discriminant validity of a measure of cognitive styles and abilities. *Journal of Personality and Social Psychology, 39*, 116-129.

Witkin, H.A., Oltman, P. K., Raskin, E., & Karp, S. A. (1971). *A manual for the Embedded Figures Tests*. Palo Alto: Consulting Psychologists Press, Inc.

An Interview With Robert Sternberg About Learning Disabilities

Robert Sternberg
Yale University

(interviewed on behalf of NAJP by)
Michael F. Shaughnessy
Eastern New Mexico University

NAJP: You recently authored a book called *Our Labeled Children* and edited a book on learning disabilities. What prompted you to venture into this realm ?

RS: Our research generally is on the topic of abilities, and of course, this category includes "disabilities" as well. In all cases, we are interested in understanding the nature of abilities and disabilities, how to identify them, and how to help all people capitalize on their strengths and compensate for or correct their weaknesses. In our book, *Our Labeled Children* (Sternberg & Grigorenko, Perseus Books, 1999), we argue that, at some level, everyone has one or more disabilities. No one is good at everything. What differs is whether the particular disability or disabilities they have are labeled by society as such. In other words, upon birth, everyone is entered into a lottery. The lottery is whether the society they are in values their abilities, and whether it recognizes their particular weakness(es) as fitting into the category of "disabilities."

NAJP: What are the difficulties in identifying children and adolescents with learning disabilities ?

RS: For reasons we describe in *Our Labeled Children*, difference scores between IQ and reading or other achievement scores are worse than worthless—they are positively misleading. As a society, we are locked into a system that is psychologically and educationally invalid and that disgraces professionals who use it. Either they know better and use it because they have to, or they don't know better, in which case the situation is even worse. In brief, the problems with difference scores pertain to several factors. First, it has been known for many years that difference scores are psychometrically unreliable. Second, IQ tests are

---

Author info: Correspondence should be sent to: Dr. Michael Shaughnessy, Psychology Dept., Eastern New Mexico U., Portales, NM 88130.

North American Journal of Psychology, 2001, Vol. 3, No. 1, 131-138.
© NAJP

inadequate as comprehensive measures of intelligence. Even psychologists who believe in the value of such tests recognize that they are only partial measures of intelligence. Third, there has been an utter failure to demonstrate that people with large differences between measured intelligence and achievement should be treated any differently from children with low achievement, regardless of IQ. In other words, the difference score is gratuitous. Fourth, the IQ tests often contain verbal material, so that some of the same variation that is in the IQ score is also in the achievement score, rendering the difference score silly because the measures are not psychologically independent. Fifth, using nonverbal IQ tests does not help, because they are even more limited in what they measure than are broader IQ tests. Sixth, difference scores lead to regression artifacts. Seventh, differences between IQ and achievement mean different things at different points along the continuum of abilities and achievement. A child could actually be labeled as having a reading disability even if his or her reading is above average, if the IQ is high enough. Finally, the populations on which IQ tests and achievement tests are normed are often different, so that the scores being subtracted are not even standardized scores relative to the same population. For these and other reasons, societal practices in the United States are inadequate. Learning disabilities should be identified on the basis of low achievement without taking into account IQ.

NAJP: What are the biological bases of reading disabilities?

RS: There is no single biological basis for reading disabilities. Differences in the *planum temporale* between children with and without reading disabilities have been identified. But the important thing to realize is that we cannot, at this point, state that these are actually "bases," that is, causal. We have established only correlations, and they do not apply to all forms of reading disabilities.

NAJP: What are the genetic bases of reading disabilities ?

RS: Many kinds of reading disabilities run in families. Differences between children with and without reading disabilities have been identified on chromosomes 1, 6, and 15. But these effects appear to be diffuse. It is unlikely that there is any one gene for reading disabilities. We live in a time when many people want to find a gene for everything. This mentality will not help us understand the complex interactions between genes and environment in the production of reading disabilities. A given genotype does not guarantee that a person will or will not have a reading disability. Genes always have their effects in interaction with the environment.

NAJP: What are the cognitive concerns and bases of reading disabilities?

RS: Most children with reading disabilities have difficulties in phonological processing of lexical information. But as we discuss in *Our Labeled Children* (and also in *Offtrack*, a book by Louise Spear-Swerling and myself published in 1996), there can be multiple cognitive signs of reading disabilities. For example, phonological deficits can lead to later problems in verbal comprehension, so that what starts off as a problem of one kind later becomes a problem of another kind. Moreover, some children have a particular problem with automatization, so that they end up processing in a controlled and often halting manner information that other children can process automatically.

NAJP: You focus in your book, *Our Labeled Children*, on reading problems. What can you add about kids with spelling learning disabilities and math disabilities?

RS: Less is known about spelling and math disabilities than about reading disabilities. Math disabilities, like reading disabilities, appear to have complex causes. There is no one "math disability." Some children have problems with computation, others with problem solving, and still others with numerical comparisons. It is important to realize as well that individuals may be falsely labeled as having a mathematical disability when in fact their problem is fear of mathematics.

NAJP: How can parents ensure that their children receive the help, accommodations and modifications that they need and deserve in the schools?

RS: If parents believe that their children are receiving inadequate services in school, they have several options. First, they can go to a psychologist to have their child assessed. This assessment may be useful if parents suspect a disability that the child's school is unwilling to acknowledge. Second, they can discuss with the school further accommodation or modification of curriculum or expectations. Third, they can work with their children at home. Fourth, they can find another school that is more willing to work with their child. Finally, in extreme cases, they can appeal to the district if they believe the school is not handling their child in an adequate way. Some parents sue, but such suits are messy, expensive, slow to conclusion, and often of little help to the children.

NAJP: What still needs to be done in terms of learning disabilities ?

RS: The main things that need to be done are the creation of better means of identification and the formulation of better curricula for helping children with learning disabilities. We are in favor of accommodations if the accommodations help children remediate their deficits. Otherwise, we believe that, in the long run, the accommodations are not always helpful. All children need to find ways of capitalizing on strengths and of compensating for or correcting weaknesses.

NAJP: Is there a best test, or I.Q. test or achievement test for identifying kids with learning disabilities?

RS: I do not think there is any one test that is an all-purpose test. For one thing, IQ tests should not even be used. A number of tests exist for identifying reading problems. It is not appropriate here, really, to promote certain tests over others.

NAJP: Do some children outgrow or grow out of learning disabilities?

RS: No, children do not outgrow learning disabilities. Children who have reading disabilities have these same disabilities as adults. What may change, however, is the individual's success in dealing with the disabilities. Often, children remediate their difficulties or develop adequate compensatory strategies. Even if the symptoms go away, however, the underlying problems do not. It is especially important to find a life path that enables us to make the most of what we do well. For example, many children with disabilities have to develop their creative and practical skills in order to learn how to adjust to their disabilities. They may be able to find a career that enables them to capitalize on these skills. Many highly successful adults were identified as having disabilities as children, or had and continue to have such disabilities, whether or not the disabilities were identified.

NAJP: What about medication? Is it helpful?

RS: Medication has been found to be helpful in some cases of attention deficit hyperactivity disorder. The decision to use medication has to be made by a physician. The long-term side effects of medications are not well understood.

NAJP: What are the current concerns in terms of learning disabilities? Where is the field going? Or are we still mired in a quagmire of the past?

RS: In education, the field continues to be mired in a political quagmire that pits tremendous vested interests on all sides against the best interests of children. When any system is adopted, certain groups of people benefit from the system and want to leave it in place. This has happened with LD, as it has happened with any other system in society. In psychology, research moves ahead at a rapid pace. Very exciting research is being done at the biological and cognitive levels. For example, psychologists are interested in pinpointing the various cognitive signs of various LDs, and biologically oriented researchers are linking various LDs to both genetic and brain functioning.

NAJP: Who is currently contributing to the field?

RS: Oh, there are many people. We tried to include some of the most distinguished contributors in our edited book, *Perspectives on Learning Disabilities* (Sternberg & Grigorenko, 1999, Westview Books).

NAJP: How much can tests tell us about learning disabled kids, how they learn and why they are having difficulty?

RS: Tests can give us diagnostic information, which can then be used in an advisory manner. Tests should be a beginning in the diagnostic and remediation processes, not an end. Skilled professional judgment always has to be brought to bear in interpreting the results of tests.

NAJP: Are there neglected learning disabilities? For example, written expression, speaking, listening, and the like?

RS: Different societies have different skills that are of particular value, and may choose to label people who are somewhat deficient in those skills as having a disability. This becomes clear if one considers the fact that in preliterate societies the concept of reading disability makes no sense. The biological or cognitive pre-dispositions may be there, but these pre-dispositions do not manifest themselves in the environment. I am weak in spatial abilities, but our society does not choose to recognize spatial skills, or the lack thereof, as a basis for a disability. However, in a hunting-gathering society, lack of spatial ability may be very compromising to adequate adaptation. Those who cannot hunt successfully literally may die.

NAJP: What roles do auditory memory, visual memory and the like play in learning disabilities?

RS: Poor memory is not generally included in the list of disabilities recognized in the United States, but as I said above, any list of disabilities is somewhat arbitrary. Poor memory is likely to lead to lower scores on conventional tests of intelligence, because such tests rely heavily on memory (e.g., of vocabulary words, or how to solve particular kinds of arithmetic problems, or of general information). Poor memory can also affect scores on tests of reading or mathematics because one has to remember specific words or arithmetic skills. This is another case where the subtraction of an achievement score from an intelligence score does not make sense: Both are dependent on memory, so that the two scores are not independent, even though they would have to be for the subtraction to make sense.

NAJP: What are the two types of approaches to understanding L.D. Why are they important?

RS: Two major approaches are sometimes referred to as intrinsic and extrinsic. The intrinsic view is that learning disabilities are largely biological and inborn. The extrinsic view is that learning disabilities are largely environmentally cause and acquired. Both view are incomplete. Learning disabilities result from interactions between biology and environment. Simplistic answers to psychological problems are attractive, but often wrong!

NAJP: What are the pros and cons of accommodations and modifications? What are the pros and cons of using technology (calculators, cassette players, spell check, and the like) for learning disabled students?

RS: Accommodations are useful to help students remediate deficits. The problem is that it has not been shown that any of the currently used accommodations benefit only students with learning disabilities. On the contrary, every accommodation currently used also benefits other students. For example, almost any student will benefit from extra time on tests. Similarly, almost any student will benefit from being able to take a test in quiet room free from distraction. So what our society has done is to create a system that gives an advantage to certain students, who may well be deserving of the advantage, without having a clear rationale for why other students are not given this advantage. This system encourages parents to go shopping for an LD diagnosis and to keep shopping until they find someone willing to give it. The system as it exists is inequitable and needs to be revised. Our goal should be to provide the best possible education for all students. The current system is not there yet.

NAJP: What are your main recommendations stemming from *Our Labeled Children* and your other edited book on Learning Disabilities?

RS: First, we need to make decisions about separate services for children with learning disabilities on the basis of matched educational and social agendas. For example, if we are going to make accommodations in the schools, then we need to make them in the workplace. Otherwise, we are preparing students for careers they may not be able adequately to fulfill. Jobs should reflect the educational opportunities that lead to them. Second, we must stop using discrepancy scores right away. Third, we should, identify children for learning disabilities on the basis of achievement alone. Fourth, we should try to prevent the development of full-blown disabilities as early as possible, rather than waiting for them to be remediated once they have fully manifested themselves. Finally, we need to help all students capitalize on strengths and to remediate or compensate for weaknesses.

NAJP: What questions have I neglected to ask?

RS: The main thing we need at this point is an integration of psychological knowledge with educational practice. Current educational practices of children with learning disabilities often reflect the state of psychological knowledge ten or even twenty years ago. Psychologists have to find a way of communicating with educators, and educators have to be ready and willing to listen.

REFERENCES

Spear-Swerling, L., & Sternberg, R. J. (Eds.) (1998). *Perspectives on learning disabilities*. Boulder, CO: Westview Press.

Sternberg, R. J., & Grigorenko, E. (1999). *Our labeled children: What every parent and teacher needs to know about learning disabilities*. Reading, MA: Perseus Books.

Robert J. Sternberg is IBM Professor of Psychology and Education in the Department of Psychology at Yale University, and Director of the Center for the Psychology of Abilities, Competencies, and Expertise at Yale. His PhD in psychology is from Stanford, and he also has received 4 honorary doctorates. He has been president of four divisions of the American Psychological Association and is editor of *Contemporary Psychology*. He has won a number of awards from organizations including the American Psychological Association, American Psychological Society, and American Educational Research Association. He is a Fellow of the American Academy of Arts and Sciences, American Association for the Advancement of Science, American Psychological Association, and American Psychological Society. He is the author of roughly 800 books and articles.

Interdisciplinarity and the Cross-Training of Clinical Psychologists: Preparing Graduates for Hybrid Careers

Daniel Holland
University of Arkansas at Little Rock

"I feel more contented when I remember that I have two professions, and not one. Medicine is my lawful wife and literature my mistress. When I am bored with one I spend the night with the other. Though this is irregular, it is not monotonous, and besides neither really loses anything through my infidelity."
Anton Chekov, 1888

The practice of clinical psychology in traditional settings may provide less career satisfaction and opportunities than in the past. Many psychology graduates have pursued specialization and sub-specialization in reaction to an unfavorable supply/demand ratio of clinical and counseling psychologists in the job market, attempting to compete with other psychologists by achieving highly focused expertise. An alternative solution is to expose psychology graduate students to interdisciplinary education and cross-training that allows them to pursue hybrid careers. Six tracks are outlined for interdisciplinary preparation and the potential development of hybrid careers: psychology-nursing/allied health professions; psychology-complementary and alternative health; psychology-administration/management; psychology-information technology; psychology-public policy; and psychology-public health. A brief description of each of the fields identified for interdisciplinary education and cross-training is presented. Each description is followed by the potential advantages to psychologists of cross-training in that field, and common prerequisites for advanced study in the field are listed. The need for significant changes in graduate education and promotion of diverse leadership in clinical psychology is discussed.

There is a growing realization that the practice of clinical psychology in traditional settings may hold less professional promise than in the past. The rapid development of managed care in medical and mental health settings has resulted in diminished rewards and liberty for many clinical psychologists. As predicted over a decade ago, economic, social, political, and internal forces have contributed to decline in the earning potential of solo private practice (Cummings & Duhl, 1987, p. 85). Furthermore, the traditional practice of psychology in medical settings

---

Author's Info: Correspondence should be sent to: Dr. Daniel Holland, Dept. of Psychology, U. of Arkansas at Little Rock, 2801 S. University Ave., Little Rock, AR 72204. dcholland@mail.valr.edu

North American Journal of Psychology, 2001, Vol. 3, No. 1, 139-170.

© NAJP

has been beset with difficulties due, in part, to the turbulence of health care in general, as well as the implicit perception of psychologists as "guests" in a domain dominated by physicians who, themselves, are experiencing significant upheaval.

Despite the constriction and downsizing that is occurring in traditional clinical settings, however, clinical psychologists retain an enormous advantage over many other disciplines. First, a distinction must be made between the profession of psychology, some areas of which are suffering, and the discipline of psychology, which continues to provide the potential for significant advantages to those educated in it. More than many other applied disciplines, education in clinical psychology, in addition to technical abilities, involves very broad concepts and skills that lend themselves to versatile application. Knowledge of such domains as behavioral health, group processes, and principles of motivation has a vast range of utility beyond the traditional mental health clinic or inpatient medical setting (Hayes, 1996). Likewise, the ability to conduct and analyze research, accomplish team goals despite differing individual perspectives, and recognize environmental as well as intra-personal contributions to behavior is a significant advantage in virtually any work setting (Hayes, 1996). This knowledge and these skills, stemming from education in the discipline, can be applied in innovative ways that extend professional opportunities and enhance career trajectories for clinical psychologists. In order to capitalize upon the potential versatility of the clinical psychology discipline, however, career tracks must be significantly diversified and new work settings must be entered. One way to achieve greater diversification of psychologists' professional roles and career opportunities is to increase interdisciplinary education and cross-training of clinical psychologists at the graduate and post-graduate levels.

The Need for Alternatives to Psychological Sub-Specialization

The profession of psychology has been undergoing two paths of development. On the one hand, the "sphere of legitimacy" for the application of psychology, in terms of what is studied, where it is practiced, and how it is delivered has expanded significantly in recent decades (Kohout, 1993; Schneider, 1981). This trend towards professional breadth and diversification is particularly evident in industrial/organizational (I/0) psychology, which has gained influence in management, education, research, and policy (Keita & Jones, 1990; Offerman & Gowing, 1990). On the other hand, another path of development, one most closely associated with clinical psychology, could be conceived as emphasizing specialties and subspecialties (Matarazzo, 1983; Mattis, 1999). This path has involved apportioning intellectual territory, distinguishing between "types" of clinical knowledge, and emphasizing partially separate identities within the profession. These separate identities, commonly identified as subspecialties within the field of clinical psychology, are built upon a meritocracy of appropriately identified post-doctoral training and subspecialty board diplomas. This professional development has particularly been advocated for clinical and counseling psychologists in medical settings (e.g. Roberts, 1998; Wiens, 1993).

There has been a continuing debate regarding how to accurately define specialties and subspecialties within psychology (Belar,1995; Krauss, Ratner, & Sales, 1997; Zilberg & Carmody, 1996). One definition set forth by the Division of Clinical Psychology (Division 12) of the APA and the Council of University Directors of Clinical Psychology (CUDCP) identified clinical psychology as a "field" (Resnick, 1991) that provides the foundation for specialties such as health psychology or neuropsychology (Wiens, 1993). These specialties, in turn, presumably, provide the foundation for newly developing subspecialties such as pediatric neuropsychology (Mattis, 1999).

It remains unclear from an empirical standpoint whether the emphasis on developing highly specialized areas of clinical psychology has resulted in improved clinical services (e.g. Faust, Guilmette, Hart, Arkes, et al, 1988) or whether this emphasis, by itself, realistically prepares students for the actual professional landscape (Prieto & Avila, 1994). It must be acknowledged, however, that some clinical specialties and subspecialties have allowed for crucial development of the field (Belar, 1997), particularly in terms of identifying how psychological principles can be applied in new and previously underserved health care settings. Furthermore, knowledge within the discipline of psychology has become vast enough that few people can master all of it, particularly for clinical application. On an intuitive level, if not an empirical one, specialized clinical knowledge and skills seem necessary for the increasingly sophisticated clinical settings in which psychologists find themselves (Belar, 1997). The question arises, then, as to how clinical psychologists will be able to pursue sophisticated specialties and sub-specialties in the field, continuing this intuitively valuable development in psychology, while remaining versatile and adaptable professionals. The answer may lie with a new model of graduate education that allows for the development of hybrid careers, so that one is able to adopt clinical specialization without sacrificing career versatility.

Re-incorporating All of Psychology's Educational Levels

In addition to an alternative to subspecialization, a new model of education is needed to address the needs of terminal masters degree graduates in clinical psychology. George Miller (1969) noted thirty years ago that "We are not physicians; the secrets of our trade need not be reserved for highly trained specialists. Psychological facts should be passed out freely to all who need and can use them" (p. 1070). His point was that it is the discipline's potential to contribute to diverse areas and interests that makes it most valuable, not its ability to cordon-off an area of practice which, as evidenced by the increasing presence of other professions in traditional psychological practice areas, is not really cordoned-off at all. Despite such admonitions, the profession of psychology has increasingly tried to protect clinical practice through gatekeeping: attempting to reserve the practice of clinical service for doctoral level clinicians by eliminating or restricting licensure at the masters level in many states (Kohout, 1993; Sleek, 1996). However, unlike medicine, which has historically succeeded in establishing sharp boundaries around its professional activities (Starr, 1982, p.225), clinical activities engaged in by psychologists are highly permeable, and other independently licensed professionals, such as social workers, nurses, or rehabilitation counselors, have successfully assumed clinical responsibilities while some of psychology's own practitioners, such as masters level clinicians, have been denied professional access (Kohout, 1993; Sleek, 1996).

The restriction and denial of professional access for masters level psychology clinicians has occurred despite there being no definitive empirical evidence suggesting masters level clinicians are less effective than doctoral level psychologists in many clinical activities (Faust et al, 1988; Goldberg, 1968; Leli & Filskov, 1981, 1984). Furthermore, this development within psychology of denying state licensure to masters level practitioners, despite the lack of empirical support for such differentiation, could be considered inconsistent with other initiatives in the field: it is occurring at the same time many doctoral level psychologists are trying to obtain prescription privileges, arguing there is clinical and empirical evidence that doctoral level psychologists are as effective in prescription skills as medically trained psychiatrists (Resnick, DeLeon, & VandenBos, 1997). What is needed at this time is a new educational model in clinical psychology, one that: 1) does not disrupt the existing scientist-practitioner model of clinical psychology education; 2) allows continued pursuit of specialties and sub-specialties without sacrificing employability; and 3) allows both masters and doctoral level graduates to benefit from education in the discipline. These three objectives can be accomplished by developing a new model of graduate education that uses the discipline of clinical psychology as one half of a professional hybrid, incorporating the knowledge and skills from other fields in order to create a truly adaptive collection of professional abilities.

Interdisciplinarity and Cross-Training: Specializing in Breadth

A recurring theme in humanities education is interdisciplinarity (Boyer, 1990, p. 21; Klein, 1990, p. 5;Kockelmans, 1979, p. 26), and in health care settings there is a growing emphasis on cross-training (Brauzer, Lefley, & Steinbook, 1996; Cordes, Rea, Rea, Vuturo, & Kligman, 1996; Moss, 1996;). Interdisciplinarity refers to the education of individuals across broad domains of knowledge, similar to the goals of a liberal arts education. Interdisciplinarity involves unifying domains of knowledge to address larger questions (Mourad, 1997), not discouraging the pursuit of specialized knowledge, but expanding the *context* of any specialization. This can be accomplished by fostering education in more than one field, preparing students of psychology to become educated in other disciplines at an advanced level. With an interdisciplinary approach to education and training, there is an emphasis upon joining together typically separate or isolated areas of knowledge for the purpose of cross-fertilization of ideas, research, and application:

> Today, interdisciplinary and integrative studies, long on the edges of academic life, are moving toward the center, responding both to new intellectual questions and to pressing human problems. As the boundaries of human knowledge are being dramatically reshaped, the academy surely must give increased attention to the scholarship of integration (Boyer, 1990, p. 21).

Interdisciplinarity also incorporates empirical efforts as well as clinical knowledge. In fact, interdisciplinary inquiry is increasingly perceived as complementary to highly specialized and focused clinical research:

> New knowledge, represented in the category of original research, is of limited usefulness if it is not integrated into a larger body of concepts and facts. For this reason, the integration or synthesis of knowledge is as valuable and as difficult as the generation of original data. A quality synthesis will reveal new patterns of meaning and can advance the field by creating new knowledge based on the integrative framework. (Halpern, Smothergill, Baker, Baum, et al, 1998).

Associated with interdisciplinarity, "cross-training" has a more technical connotation and refers to the acquisition of specific skills garnered from separate disciplines, such as learning both nursing skills and psychotherapy skills, or psychological consultation skills and personnel management, rather than specializing in just one of these areas. The emphasis on interdisciplinarity and cross-training is not necessarily counter to the emphasis on specialization, since it does not preclude specialization as one component of a broad, interdisciplinary education:

specialized knowledge can still be sought, but it can be sought in conjunction with education and training in other disciplines and skills.

The Task Force of the Society for the Teaching of Psychology recently proposed changes to the very manner in which scholarship in psychology is conceived and judged, calling for a broader view of what constitute scholarly activities, consistent with Boyer's (1990) proposals in his seminal book, *Scholarship Reconsidered* (Halpern et al, 1998). Interdisciplinary innovations within the larger academic community are already occurring. For example, an arrangement at Tulane University called the Interdisciplinary Scholars Network focuses on creating an interdisciplinary network of graduate and professional students who can interact and share knowledge through organized meetings and a lecture series. Students participating in this network do not sacrifice their areas of specialized interest, but have the opportunity to reframe these interests in an interdisciplinary context among students from related and distant disciplines. Within the discipline of psychology, the psychology department of Seattle Central Community College developed a Coordinated Studies Program based on a collaborative-learning model that links the discipline of psychology to multiple other disciplines in order to promote breadth of knowledge (Finley, 1995). Similarly, Industrial/Organizational (I/O) psychology in general has been successful in achieving an interdisciplinary educational mission while preserving its own identity as a field within the broader discipline of psychology (Arthur & Benjamin, 1999, p. 100). These innovative and adaptive approaches to education herald a new era in preparing students for the world of work. In a time of rapid change in health and mental health settings (Broskowski, Marks, & Budman, 1994,p.1; Wheatley, DeJong, & Sutton, 1997), interdisciplinarity and cross-training may represent one of the most adaptive and constructive approaches to the education of clinical psychologists in graduate programs.

A Proposal for Hybrid Careers in Psychology

There has been a growing and persistent call for clinical psychologists to become more familiar with other disciplines in order to become more versatile and extend the influence of the profession more broadly (i.e. Bray, 1996; Cox, 1997; Haley et al, 1998; Kohout, 1993; Wiggins, 1994;Yung, Hamond, Sampon, & Warfield, 1998). There have been few accounts, however, of actual interdisciplinary or cross-training models linking psychology to other disciplines at the graduate level. Notable examples of graduate interdisciplinary education in psychology are the psychology and law programs, such as those at the University of Nebraska's interdisciplinary graduate program in which students can gain a Ph.D. or masters in psychology as well as a law degree, or the interdisciplinary programs in law and psychology at Simon Fraser University in British Columbia (Ogloff, 1999). Despite the greater versitility of psychology as a discipline, medical education has actually more aggressively pursued interdisciplinary educational options: many medical schools have offered interdisciplinary and dual degree programs for many years (e.g. Raymond, 1999). Some innovative proposals and published anecdotal examples of interdisciplinarity and cross training in psychology include: interdisciplinary study of psychology and government policy (Walley, 1995); uniting psychology with nursing (Thomas, 1996); psychologists learning to conduct physical examinations (Folen, Kellar, James, Porter, & Peterson, 1998); and training psychologists in psychopharmacology for the purpose of prescription privileges (i.e. Brentar & McNamara, 1991).

Rather than pursuit of highly specialized clinical post-doctoral training as the *only* form of post-graduate professional education encouraged in clinical psychology, it is proposed here that interdisciplinarity and cross-training in other fields, in lieu of subspecialty post-doctoral residencies for some individuals, can potentially serve the profession of clinical psychology, its graduates, and the public very effectively. The cross-trained psychologist would be able to enter more than one professional arena, bringing with her the knowledge and professional title of a clinical psychologist as she influences and shapes other fields that are usually beyond the direct influence of most psychologists. Equally important, her own career options would be enhanced during a time of increasing constraint in many clinical psychology roles.

If clinical psychology graduate students were exposed to coordinated interdisciplinarity as part of their graduate curriculum, they could subsequently pursue cross-training in another field upon completion of their degree program in psychology, thereby receiving what would functionally be post-graduate training in another independent profession and achieving a hybrid education. The psychologist-public health specialist, psychologist-nurse practitioner, psychologist-massage therapist, or psychologist-business manager would be in a position to pursue a greater diversity of professional opportunities and avoid temporary recessions in any one field. Furthermore, like the psychologist-attorney educated in one of the M.A./ J.D. or Ph.D./J.D. programs, these additional hybrid professionals would bring a synergistic knowledge to both fields and to the public.

In order to be prepared to enter advanced education in another field upon completion of masters or doctoral education in clinical psychology, graduate programs facilitating this cross-training model would need to incorporate the other discipline's prerequisites into the graduate psychology curriculum. For example, a graduate psychology student might be taking courses in computer programming, management information systems, and engineering, in addition to psychology core courses, in order to qualify, upon graduation, for advanced post-doctoral study of a technological field like Information Technology.

What follows are brief outlines of six broad hybrid paths for interdisciplinary learning and cross-training, linking clinical psychology to other professional fields. The six hybrid paths outlined are: psychology-nursing/allied health professions; psychology-complementary and alternative health; psychology-administration; psychology-information technology; psychology-public policy; and psychology-public health. The psychology-law hybrid has not been included here, since that particular model has been covered extensively elsewhere (e.g. Bersoff, Goodman-Delahunty, Grisso, et al, 1997). Likewise, the cross-training necessary for psychologists to obtain prescription privileges is not included here, as that model has been given extensive consideration previously (e.g. DeLeon, 1993; Resnick, DeLeon, & VandenBos, 1997) and, until laws are changed, does not presently constitute a career option for psychologists.

The emphasis of each of the six outlines is on: 1) a brief description of the field proposed for cross-training; 2) the advantages to psychologists of pursuing cross-training in that specific field; and 3) the prerequisites or recommended course work that would have to be incorporated into the graduate psychology curriculum in order to facilitate post-graduate cross-training in the outside field. Such prerequisites or recommended course work (beginning accounting, for example, for those students wishing to prepare for both psychology and business; or anatomy and physiology for those wishing to pursue both psychology and nursing) would likely not be gained within the psychology department, but would be fulfilled by students engaging in interdisciplinary education by taking such courses in other departments of the university. The model, therefore, is one of interdisciplinarity at the graduate psychology level, with advanced cross-training and education in another field following completion of the graduate psychology degree. At the completion of such an interdisciplinary and cross-training program, many graduates would hold dual degrees, such as a masters or Ph.D. in psychology and a nursing degree or masters in business administration. *All* graduates, though, with or without dual degrees, would have achieved an extensive interdisciplinary graduate education, with hybrid skills and advanced knowledge of more than one field.

SIX HYBRID PATHS FOR THE EDUCATION OF CLINICAL PSYCHOLOGISTS

Psychology and Nursing/Allied Health Professions

Allied health professions have traditionally been all of those in which the clinicians assist, facilitate, or complement the work of physicians. In the past, such professions typically included nursing, physical therapy, occupational therapy, speech language pathology, audiology, physicians assistants, and others. The focus here is on nursing, occupational therapy (OT), and physical therapy (PT), though conceivably many of the disciplines comprising allied health could provide opportunities for interdisciplinary education and cross-training of clinical psychologists. The current growing independence of many of these allied health professions from physician dominance is now resulting in a need to reconsider the term "allied health" entirely. Advanced practice nurses (APNs), for example, are gaining independent prescription privileges on par with medical personnel (Segal-Isaacson, 1998), and OTs and PTs are actively ad- dressing how to monitor and manage their own clinical treatment hierarchies over which they have significant control (Holland, 1998). The former College of Health Related Professions at the University of Florida in Gainesville changed its designation to the College of Health Professions in 1995 to more accurately reflect the role of disciplines such as OT, PT, and SLP ("Changing Name", 1995). What follows are specific definitions and descriptions of the three health professions proposed here for cross-training psychologists: nursing, OT, and PT.

Nursing

There are multiple levels of nurses. The Registered Nurse (RN) is the recognized standard of practice and the RN constitutes the most common educational level for nursing. There are generally four routes to becoming an RN: 1) a diploma program offered directly through a hospital (approximately two to three years following high school); 2) an associates degree through a junior college (a two year associates degree following high school); 3) a baccalaureate nursing degree through a four year college or university program which frequently results in either a bachelors degree in nursing (BN) or both a non-nursing bachelors degree accompanied with a nursing diploma; and 4) a *pre*-nursing background within a baccalaureate program followed by a post-baccalaureate diploma program or two year RN program.

Education beyond the RN or BN degree, usually a masters degree, frequently results in one of four types of advanced practice nurse (APN): nurse practitioner (NP), clinical nurse specialist (CNS), nurse midwife (NM) , and nurse anesthetist (NA; Pearson, 1999). While the four types of advanced practice nurses frequently differ with regard to practice specialties, practice settings, and emphases, what they share, in addition to a foundation in the nursing philosophy, is rapid advancement and growing influence in the contemporary health care environment despite the reticence and objections of organized physicians' groups (Cooper, Laud, & Dietrick, 1998).

Occupational Therapy

Occupational therapists work to restore the highest possible level of independence of those individuals limited by physical injury, illness, cognitive impairment, psychosocial function, or developmental or learning disability (Neistadt & Crepeau, 1998, p. 897). Occupational therapy does not focus exclusively on physical function or rehabilitation, but rather the development of whole activities of daily living. Occupational therapists, therefore, address the physical and cognitive aspects of recovery or adjustment.

Occupational therapists are educated at either the bachelors or masters level. Occupational therapy is recognized as a crucial discipline within medical and rehabilitation settings, though it is often not well understood by the public. Occupational therapy is rapidly moving into areas of clinical treatment and program administration that significantly impact and overlap with clinical psychology in medical and mental health settings.

Physical Therapy

Physical therapists (PTs) evaluate and treat the physical capacities and limitations of individuals with illness or injury (Neistadt & Crepeau, 1998, p. 793). Physical therapists are often primarily focused on patients' ability to increase strength, balance, and mobility (APTA, 1995). Unlike OT, PT tends to focus more on physical functioning and recovery rather than broad activities of daily living.

The Benefits to Psychologists of Cross-Training in Nursing/Allied Health Professions

The mutually beneficial alliance between the fields of psychology and nursing has been emphasized in the past (DeLeon, Kjervik, et al, 1985), and the advantages of being cross-trained in clinical health psychology and nursing have been outlined (Thomas, 1996). Increasingly, greater responsibility for patient care is being given to RNs and APNs in hospital and clinic settings in order to reduce health care costs (Brink, 1999). The result is higher profiles and greater potential clout for the nursing field within managed care settings.

Median starting salaries for some APNs now frequently exceed that for clinical psychologists in many markets (Fisher 1997, p. 481). NPs have been particularly successful in lobbying for prescription privileges in a large number of states (Sharp, 1999). Clinical psychologists educated in and holding a license as an NP would likely be in an optimal position to assume highly sophisticated and dynamic positions in a wide variety of clinical settings. Additionally, the National Institute for Nursing Research (NINR), within the National Institutes of Health (NIH), is already conducting some areas of research focusing on psychological variables (Sheridan, 1999). Cross-trained clinical psychologist-nurses would be well positioned to contribute to and lead such empirical efforts in a wide variety of health care and research settings.

While the health professions of occupational therapy and physical therapy have recently suffered constricting job markets and slowed employment and financial growth due to reformed Medicare reimbursement associated with the Balanced Budget Act of 1997 (Schaffer, 1998), future projections for these fields are very positive (U.S. Department of Labor, Bureau of Labor Statistics, 1998, p. 194). The United States demographic composition is a harbinger of heightened need for such rehabilitation and geriatric care professionals in the future (VandenBos & DeLeon, 1998, p. 19). Clinical psychologists cross-trained in one of these other health care fields would be particularly well suited for clinical-administrative roles and team leadership in interdisciplinary rehabilitation settings. Clinical psychologists possessing hybrid skills and dual licenses would also have the advantage of expanded employment opportunities and versatility in a volatile health care job market.

The fields of clinical psychology must begin to consider that current needs in health care might best be served by advanced, cross-trained professionals with multiple clinical capabilities. Psychologists, in order to continue developing influence in health care settings, will need to incorporate a much broader repertoire of clinical skills and functions (Holland, 1998; Troy & Shueman, 1996, p. 67). Ideally, clinical psychologists in medical settings will begin receiving education in the clinical knowledge and skills of other successful health care disciplines, achieving a more holistic function as providers. The psychologist-nurse practitioner, psychologist-OT, or psychologist-PT, for example, would be in a posi- tion to 1) address both physical and psychological needs in an integrated manner; 2) take advantage of career opportunities in either profession; and 3) conduct particularly sophisticated interdisciplinary research.

Clinical psychology need not wait only for legislation that will allow for prescription privileges for psychologists in order to expand their health care role. Multiple opportunities for expansion of clinical health care abilities are currently available for those psychologists willing and able to pursue post-graduate education in one of the closely related, but independent, health professions. Furthermore, as noted above, some health professions, like nurse practitioners, have already gained prescription privileges. Interdisciplinary education that allows students to subsequently pursue post-graduate training in one of the allied health fields may facilitate psychologists gaining prescription privileges through the legislative progress made by the *other* profession.

Frequent Prerequisites for Nursing/ Allied Health Professions Programs

Each of the health fields reviewed above has its own set of prerequisites or recommended courses prior to advanced study. Some of the health professions, such as Advanced Practice Nursing, have prerequisites that would necessitate a great deal of navigation in order to integrate them into a graduate clinical psychology curriculum. In the case of any of the Advanced Practice Nursing specialties, for example, one must already possess an RN degree in order to pursue a masters degree. The interdisciplinary preparation in a graduate psychology program would ideally involve course work and clinical rotations that would result in an RN degree as well as a masters or Ph.D. in clinical psychology. Graduates of such an interdisciplinary psychology-nursing program could subsequently go on to pursue advanced study in nursing resulting in Advanced Practice Nursing status and the contemporary liberties that come with it.

Fulfilling the prerequisites for advanced study of OT or PT would likely be less strategically challenging than for nursing, since the traditional graduate curriculum in clinical psychology often significantly overlaps with common prerequisites for advanced study of OT and PT. For OT, many graduate programs require both social science and natural science didactic background. Common social science prerequisites include Abnormal Psychology, Developmental Psychology, and Sociology. Natural Science prerequisites include Physiology, General Anatomy, Physics, Chemistry, Introductory Statistics, and computer literacy.

Common PT prerequisites include those listed above for OT, but with frequent additional prerequisites of microbiology and exercise physiology. Many PT programs do not emphasize a substantial social science background that most OT programs prefer.

Psychology and Complementary and Alternative Health

Complementary and alternative health (CAH) refers to a range of healing and wellness methods that complement conventional biomedicine and broaden health care options (Berger, Hawley, & Pincus, 1992). Many CAH practices have been drawn from traditional Eastern systems of health (such as Indian Ayurvedic approaches to health, Chinese medicine, and Tibetan medicine) and native Western approaches to health (such as Native American healing practices). CAH tends to have three broad areas of focus: biophysical modalities (e.g. acupuncture, Alexander Technique, massage, subtle energy, yoga, meditation, qiqong) biochemical modalities (e.g. herbs, nutrition, and some aspects of acupuncture) (Wu, Bandilla, Ciccone, Yang, Cheng, Carner, Wu, & Shen, 1999), and spiritual practices (e.g. meditation, yoga, prayer, pilgrimage). The emphasis of most complementary and alternative health is holistic, so there is a great deal of intentional integration and overlap among these three broad modality domains.

The Benefits to Psychologists of Cross-Training in Complementary and Alternative Health

Complementary and alternative health (CAH) has experienced dramatic growth due to increased demand over the past decade (Eisenberg, Kessler, Foster, Norlock, Calkins, & Delbanco, 1993; Rosellini, 1999). The Office of Alternative Medicine (OAM) of the National Institutes of Health (NIH) was recently elevated to a free-standing center within the NIH (Muscat, 1999). Now called the National Center for Complementary and Alternative Medicine (NCCAM), the center is now able to fund its own research grants independent of other NIH institutes or centers (Muscat, 1999). Over the past five years, new empirical peer-reviewed journals, such as Alternative Therapies in Health and Medicine, focusing exclusively on CAH, have emerged. International, interdisciplinary conferences in topics like traditional Eastern medicine have also begun to take place.

CAH may represent one of the most promising interdisciplinary and cross-training opportunities for contemporary clinical psychologists, though little has been proposed regarding how psychology might interface with CAH in terms of educational preparation or practice. CAH is projected to continue to grow rapidly (Boucher & Lenz, 1998) and certain practices within CAH like massage therapy (Thomas, 1999), acupuncture, and naturopathic medicine (Cassileth, 1998, p. 52) are increasingly being incorporated into mainstream Western medical practice. A number of medical schools have begun to incorporate CAH into the curriculum (Rosellini, 1999). While clinical psychology and other non-physician health care professionals have experienced resistance from the American medical establishment with regards to prescription privileges and other avenues for growth (Klusman, 1998; Starr, 1982, p. 27), such resistance does not exist in areas of complementary and alternative health. In fact, a number of CAH leaders have called for increased collaboration with and inclusion of clinical psychology (Horrigan, 1998; Menkin, 1999). It is also worth noting that, at a time when psychology is experiencing increased difficulty with reimbursement for clinical services, CAH is slowly beginning to experience increased coverage and inclusion (Goleman & Gurin, 1993, p. 17; Cassileth, 1998, p. 52). Since many principles and practices being attributed to CAH are, in fact, traditional psychological concepts and methods (relaxation training, meditation, biofeedback, hypnosis, etc.), the opportunity for clinical psychologists to contribute to and benefit from involvement in the fields comprising CAH is already evident. Some CAH practices have been criticized on the grounds that they lack empirical support (McCutcheon, 1994; 1995) or have not yet been subjected to randomized clinical trials (Angell & Kassirer, 1998). Clinical psychologists cross-trained in CAH could bring research skills and a demanding empirical perspective to CAH that could help strengthen the field. As health care professionals, it could be said that psychologists are obligated to become more involved in CAH, particularly given the public's growing pursuit of services subsumed under this designation, and the need for empirical influence. It is notable that the public has been very vocal in its pursuit of CAH and has advocated for the inclusion of CAH in health maintenance organizations (HMOs) and insurance coverage (Moore, 1997). The field of clinical psychology, by becoming more closely involved with CAH, may be able to better identify and incorporate those qualities possessed by CAH and its practitioners that have resulted in such strong public appeal.

Frequent Prerequisites For Complementary and Alternative Health Programs

The array of complementary and alternative health practices is so broad that identifying prerequisites for entry into educational or training programs is somewhat arbitrary, particularly since some training opportunities in modalities like meditation or yoga are not academic and do not have prerequisites. There are a number of formal educational opportunities in complementary and alternative medicine, however, that do have specific course requirements. A large number of free standing institutes of acupuncture and oriental medicine offer degree or certificate programs in oriental medicine or acupuncture. Many such programs have a small number of prerequisites highly relevant to graduate education in CAH. Even when training programs do not impose specific academic prerequisites, knowledge of human biology, anatomy, and physiology would clearly provide useful foundation knowledge before entering any rigorous training program in holistic health, particularly since such training programs are not likely to be remedial and may assume such background knowledge. Therefore, clinical psychologists pursuing CAH cross-training opportunities that do not have prerequisite academic courses could still clearly benefit from interdisciplinary graduate psychology education that incorporates the courses listed below. These courses are drawn from some of the most common prerequisites or suggested courses for programs in Oriental Medicine and complementary health (American Association of Oriental Medicine, 1998) and Naturopathic Medicine: Anatomy and Physiology; Pharmacology; Comparative Religion/Eastern Spirituality; Human Biology; Organic Chemistry; General Chemistry; and Physics.

Psychology and Public Health

Public health is a social institution, a discipline, and a practice (Millbank Memorial Fund Commission as cited in Last, 1995, p. 134). Public health involves community-centered endeavors, the goals of which are to protect the community against threats to its health (Fairbanks & Wiese, 1998, p. 5). The discipline of public health, like the concept of health itself, has evolved towards a "process" definition (Fairbanks & Wiese, 1998, p. 4), involving efforts across health care, science, and social science disciplines, government agencies, and national boundaries. The most common setting for public health professionals is still in government agencies, though the need for public health knowledge in managed health systems and community prevention programs is well known (Fairbanks & Wiese, 1998, p. 28).

The Benefits to Psychologists of Cross-Training in Public Health

A decade ago, it was noted that there were few clinical psychologists working on macro-level public health issues (Winett, King, & Altman, 1989, p. 5). The rapid growth of clinical health psychology (Belar, 1997; Sheridan, 1999), however, has resulted in a great deal of overlap between public health and the discipline of psychology. Clinical health psychology is often closely allied to the field of public health in its mission (Koop, 1983), and the scientific knowledge base deemed important to clinical health psychology overlaps much of what is central to the field of public health (Sarafino, 1998, p. 19).

A number of proponents for the specialty of clinical health psychology have advocated for interdisciplinary education in both psychology and public health (Winett, King, & Altman, 1989, p. 26). Given the current clinical and financial incentives for prevention in the contemporary managed health care environment, an interdisciplinary background in these two areas would fill a growing need (Lasker & The Committee on Medicine and Public Health, 1997, p. 44). There are a number of dual degree MA/MPH or Ph.D./MPH programs available, uniting psychology and public health; the potential to join the two disciplines, however, has yet to be fully realized.

There is a growing call for the expansion of public health knowledge and application worldwide (World Bank, 1993, chap. 7). This knowledge and these skills are likely attainable by clinical psychologists without additional formal public health didactics or an additional degree program, though formalized education such as that in an MPH degree program would likely provide the greatest professional validity and recognition for those psychologists who wish to pursue public health endeavors. Furthermore, as a health profession, clinical psychology has not pursued international service the way the fields of medicine, nursing, and, particularly, public health have. The clinical psychologist-public health professional could more readily enter into causes like international health that might not be as accessible to a traditionally educated clinical psychologist.

Professional success will be achieved by those who possess valuable knowledge and combinations of skills that are in greater demand than supply. Those who unite knowledge of both clinical psychology and public health, similar to physicians who have sought public health knowledge, will likely represent valued and unique professionals able to address health concerns, policy, and prevention efforts in ways others cannot (Lasker & The Committee on Medicine and Public Health, 1997, p. 36).

Frequent Prerequisites for Graduate Study in Public Health

Many graduate programs in public health prefer that applicants have a background in both natural sciences and social sciences, consistent with the discipline's focus upon both the biological and social factors underlying health risks. Natural science courses including biology, physiology, and microbiology; and social science courses including sociology, anthropology, and economics, in addition to psychology, are frequently identified as providing important background knowledge for the graduate study of public health.

Psychology and Administration/Management

Peter Drucker (1990) claims the job of a good leader is not simply to capitalize on charisma, but to formulate problems and organize others in the pursuit of a mission, and to rally and lead others in the pursuit of well defined goals (p. 3). The most important skills needed for administration or management are often debated, but there are some core skills and fundamental areas of knowledge that consensus suggests do, in fact, inform administrative functions. These areas of knowledge and skill frequently involve understanding personnel issues, organizational behavior, finance, operations, organizational strategy, accounting, quantitative analysis, and marketing (Silbiger, 1993,p. 5). Some of these skills are closely associated with those learned in clinical programs (personnel issues, for example), while others are further afield (e.g. accounting). It is, however, such partially overlapping professional domains that may offer some of the most promising hybrid career opportunities and are the focus here.

The Benefits to Psychologists of Cross-Training in Administrative/ Management Skills

Administrative and management skills are not currently a core component of the graduate psychology curriculum in most clinical programs (Cox, 1997; Maddi, 1997), though a number of authors in a number of health care fields have emphasized that clinicians could benefit significantly from greater exposure to such skills, preparing for positions as clinician-managers in a variety of settings (Cordes, Rea, Rea, & Vuturo, 1996; Holland, 1998). A survey of psychiatrists found that those in health care administrative positions reported performing a greater variety of tasks and experiencing greater job satisfaction than staff psychiatrists (Ranz & Stueve, 1998). Learning such administrative and management skills can be accomplished in a number of different ways and through a number of different disciplines. Exposure to I/O psychology at the graduate level is one of the most efficient and readily achieved means of gaining such skills for most clinical psychology students in departments that also offer an I/O specialty. Business school, public administration programs, and health management programs within schools of public health also offer opportunities for the graduate or post-graduate clinical psychologist looking to cross-train in administrative skills. The hybrid clinical psychologist-manager would conceivably have two career paths from which to choose: one would constitute entering traditional health or mental health settings as an administrator who also possesses first hand knowledge of clinical service provision; the alternative would be to pursue a career outside traditional clinical settings entirely, in an area such as corporate management or consulting, where a clinical background and knowledge of psychology would broadly inform one's work activities, but not determine one's primary professional identity.

The professional educated in both clinical psychology and management would be in a position to manage both the financial challenges of an organization and personnel challenges in a highly sophisticated manner. Given the professional success of I/O psychologists (Erffmeyer & Mendel, 1990; Schippman, Hawthorne, Schmitt, 1992) clinical psychology could likely gain significantly by integrating management skills into the graduate curriculum and preparing some students for post-graduate pursuit of management degrees.

Frequent Prerequisites for Administrative/Management Programs

There are a number of different genres of graduate administration and management education, each with its own emphasis and, as a result, its own set of common prerequisites or suggested background knowledge. Many administration and management programs do not have specific course prerequisites for admission, but clearly some background knowledge would be advisable, if for no other reason than to orient the student to the kind of management program she might prefer. Administrative skills gained through a masters of public health (MPH), for example, are likely to be those needed for health services management, relief organizations, or Federal health agencies, and warrant background in the biological and quantitative sciences that are likely already fulfilled through the curriculum of most traditional clinical programs. Administrative skills gained through a masters of business administration (MBA) program, on the other hand, are likely to have a corporate emphasis, a greater focus on finance and e-commerce, and require competencies in areas like accounting and budgeting that are not common content in most clinical programs. For those pursuing administrative and management skills through post-graduate study in masters of public administration (MPA) programs, which are primarily concerned with leadership in the non-profit sector (McCormick, 1994, p. 1), background coursework in accounting, finance, government policy, and legislative process might prove helpful. Below are common prerequisites for graduate programs in business administration and public administration. Although many schools of public health also provide programs in health management, prerequisites for entry into public health programs have been covered above.

There are an enormous number of MBA programs and no universal set of academic prerequisites. Some programs, however, do either require or strongly favor applicants with two to three years of work experience. Some of the additional common academic prerequisites or suggested competencies include: College Algebra, Integral and Dif- ferential Calculus, and Basic Accounting. Since MBA programs tend to emphasize economics, finance, accounting, and marketing (in addition to statistics and organizational behavior which most psychology graduate students would know well), students possessing familiarity with some or all of these areas would make a ready transition into a business curriculum. Furthermore, some business schools offer accelerated degree programs for incoming students possessing significant background coursework in business. Graduate clinical psychology programs could readily integrate business school prep courses into the traditional curriculum, and some programs could likely achieve a dual degree program in clinical psychology and business that would significantly enhance the career options of graduates.

MPA programs offer education for the management and administration of public and non-profit programs (McCormick, 1994, p. 1). They therefore tend to emphasize knowledge of the social sciences, much of which a graduate program in psychology is likely to satisfy. Suggested, though frequently not required, background courses for admission to graduate public administration programs include: Economics, Anthropology, Political Science, and Statistics.

Psychology and Information Technology

Information technology is a term that refers to a broad spectrum of functions and applications, ranging from record-keeping activities to the interaction of human factors with computers. In the most practical sense, information technology refers to computer applications aimed at capturing, transmitting, and recording information for use by an entire organization (Griffith, 1995, p. 418). Information systems are often networked in complex ways, and this complexity requires specialized knowledge and skills in order to engineer, organize, and integrate these systems within organizations. For example, an integrated system of information in a health care setting might involve combining patient accounting and medical records (Griffith, 1995, p. 419), two distinct functional units that must continuously share information. Those who engineer and construct these integrated networks of information constitute critical resources within an organization dependent upon such coordinated processes.

The interaction of human factors with computer science is also subsumed under information technology and incorporates a great deal of psychological and cognitive principles along with those of information technology. Similar to the way in which ergonomics addresses the physical relationships between people and machines (Barlow, 1990, p. 59; Quick, 1999), parallel concerns now focus on the interaction between human cognitive factors and computer technology. This domain of information technology is interdisciplinary in nature, frequently concerned with identifying the most productive and "user friendly" applications of technology, and offers promising opportunities for students with advanced knowledge of clinical psychology.

The Benefits to Psychologists of Cross-Training in Information Technology

There is an increasing need for clinical psychologists to become more engaged with technology and computer science (Berry, 1996). While other specialties within psychology have sought further collaboration with technological fields and computer science, giving rise to new and innovative hybrid disciplines, clinical psychology has remained less technologically invested. The area of artificial intelligence, for example, is a discipline entrenched in both cognitive psychology and technology (Wagman, 1998, p. 12), though the contribution of applied clinical knowledge to this area has yet to receive full attention. Nevertheless, the involvement and productivity of other fields of psychology in technology and computer science suggests promising opportunities for clinical psychologists as well. The School of Information at the University of Michigan, for example, has a specialization in Human-Computer Interaction which specifically emphasizes organization theory and behavioral science along with computer science, and a number of faculty in this specialization have backgrounds in cognitive psychology. A number of other universities have programs in information technology and computer science that are highly accessible for interdisciplinary education. The University of Arkansas at Little Rock is in the process of developing a new College of Information Science and Systems Engineering that, in addition to an academic major, will offer an academic minor in information systems and will encourage students with a variety of career interests to adopt this minor.

Clearly, the potential for psychology in general to contribute to the field of information technology has begun to be realized. The opportunity for clinical psychology specifically to become involved also seems very promising. For example, graduates with a knowledge of both behavioral health and information technology would be well positioned to help design or maintain management information systems (MIS) in managed care or health insurance settings. Unified knowledge of clinical psychology and information technology would also result in hybrid professionals capable of designing distance learning methods and computer software that educates the public, students, or professionals in matters related to behavioral health. Psychologists at the Missouri Institute of Mental Health, for example, designed an educational software program that introduces the lay public to mental health and behavioral health concepts in a very accessible and user-friendly format (Epstein, Sage, & Wedding, 1995). Furthermore, clinical psychology graduates who are crossed-trained in information technology might be best positioned to help design and influence the production of computer-based approaches to psychological assessment and treatment, including telehealth technology (Stamm, 1998), computer administered cognitive testing, and computer-generated reports.

Frequent Prerequisites for Graduate or Advanced Study of Information Systems

Programs of study in information technology are frequently free of uniform prerequisites, though background experience or course work in some areas is advised. For example, the graduate degree program in the School of Information at the University of Michigan does not have strict prerequisites, but reviews applicants on a case-by-case basis for fit with the goals and demands of the program. A clinical psychology student applying to the program would be considered most competitive if he had: 1) a psychology background that included cognitive courses; 2) a foundation in technology, preferably with some programming experience; and 3) an interdisciplinary background that might include anthropology, economics, library or information science, architecture, sociology, engineering, and/or mathematics (G. Furnas, personal communication, March 30, 1999). Furthermore, a formal degree program in Information Technology may not be necessary for many psychology graduate students wishing to unite clinical psychology with Information Technology skills, though a very strong foundation of knowledge in technology, however acquired, is critical.

Psychology and Public Policy

Public policy is ultimately that for which the government exists (Davidson & Oleszek, 1998, p. 349). It is what the government says and does about perceived problems (Ripley & Franklin, 1991, p. 12). Public policies may be explicit, authoritative statements of what the government is doing about a specific issue, written down in the form of laws or rules (e.g. financing social security or naming a federal building); or they may be far less formal (e.g. a stated mission in the President's State of the Union Address) (Davidson & Oleszek, 1998, p. 350). Policies may emphasize substance (e.g. building the defense system or highway infrastructure) or procedure (e.g. mandating program management standards) (Davidson & Oleszek, 1998, p. 350). Policies emerge through a Darwinian-like process of selection, determined by a host of factors and inputs. Understanding the complex process of policy formation can be difficult, but such an understanding is imperative for professional psychologists, if they are to successfully contribute to the shaping of major policy decisions.

The Benefits to Psychologists of Cross-Training in Public Policy

Professional psychology has a long history of interest in social issues and public policy. Social psychology and community psychology are fields that clearly represent this investment. In 1969 the theme of the APA Annual Convention was "Psychology and the Problems of Society" (Korten, Cook, & Lacey, 1970, p. xi). Despite this consistent interest, however, few clinical psychologists have been actively involved in policy making (DeLeon, 1988). It has been stated with increasing urgency that psychologists must begin to realize the role they can have in the process of making public policy if they are to influence legislation and political actions (DeLeon, 1988; Lee, DeLeon, Wedding, & Nordal, 1994).

One way to foster greater involvement of clinical psychologists in the public policy sphere is to encourage interdisciplinary education at the graduate or post-graduate level that specifically emphasizes public policy, government, and the legislative process. Graduate programs in clinical psychology have not typically encouraged development of high level public service professionals (DeLeon, 1988). Partially correcting for this situation, since 1974 the APA has promoted and financed a congressional fellowship program in order to facilitate the involvement of psychologists in the legislative and policy-making process (Lee et al, 1994), which provides an opportunity for psychologists to interface with the political world. This APA fellowship program is small, however, and only a handful of applicants are chosen for the opportunity each year. At this critical point in the development of the profession, when policy makers are directly influencing the day-to-day functions and livelihoods of clinical psychologists through decisions affecting health care finance, mental health access, patients' rights, children's health, and health research funding, the time has clearly arrived for education in clinical psychology to promote advanced knowledge of the legislative process and public policy on a larger scale.

A number of options exist for clinical psychologists to learn a great deal about the legislative process and public policy. One option would be the equivalent of "executive education" in legislative and policy studies. Such options are available through programs like that offered by the Government Affairs Institute at Georgetown University, where individuals can earn a certificate in legislative studies by attending intermittent weekend and week long intensives on Capitol Hill. Other options include continuing education through local universities or state government agencies, where courses and seminars on government policy and the legislative process are frequently open to those who are interested.

The most intensive option would involve interdisciplinary and/or post-graduate study of public policy, possibly resulting in dual degrees in psychology and a masters in public policy (MPP), but more importantly resulting in a thorough understanding of the workings and functions of government and legislation. Regardless of the specific means of gaining education in public policy, however, what currently remains essential is for graduate education in clinical psychology to foster accessibility to public policy knowledge, since without such knowledge the field as a whole will have a diminished voice on the national scale.

Frequent Prerequisites for Graduate Study in Public Policy

Graduate programs in public policy, like the development of public policy itself, draw from a diverse collection of educational and professional backgrounds. Many masters of public policy (MPP) programs prefer a history of employment or volunteer work that involved some aspect of policy or legislative knowledge. Most programs assume an understanding of the structure of American government and the legislative process. Specific didactics that are frequently required or preferred for the advanced study of public policy include micro- and macro-economics, American history and political science, and college algebra or statistics.

Discussion

The proposal to begin developing hybrid career tracks for clinical psychologists through interdisciplinary graduate education and post-graduate cross-training is likely to generate controversy. Such a model of graduate education necessarily requires a certain relaxation of professional identity, a willingness to advocate for collections of skills, rather than membership in a profession per se, as a new approach to a career for some graduates. This represents a radical departure from the historical emphasis of clinical psychology, which adopted and perpetuated a guild approach to professional identity (Kiesler, 1987, p. 57), emphasizing adherence to the specific and traditional roles of a clinician-researcher in order to maintain professional salience. Any controversy raised by the proposal to educate graduates for hybrid careers, however, would be productive at this time, since the need for dramatic change in graduate psychology education has clearly arrived, and dialogue is needed.

If graduate study in clinical psychology continues to result only in job choices among traditional clinician-researcher roles, clinical psychology programs will lose talented students to other graduate disciplines and professional programs that promise graduates clear career trajectories. The job market for traditionally educated and trained clinical psychologists is no longer a selling point for the discipline. What graduate education in clinical psychology must now do is redesign graduate programs so that students can continue to gain the knowledge and skills of the discipline, but have career opportunities than transcend the traditional professional boundaries.

While debate and argument are healthy for any large profession, there is at this time a significant amount of divisiveness within clinical psychology. This divisiveness is largely a result of an oversupply of traditionally educated clinical psychologists in a limited scientist-practitioner job market. The oversupply of clinical psychologists results in a reduction of professional and personal rewards due to competition for shrinking resources (Robiner, 1991; Wiggins, 1994). As a result of this unfavorable supply/demand ratio, members of the profession are engaging in turf battles over traditional practice areas, attempting to claim certain clinical activities or clinical populations for subspecialties designated by board certification (e.g. Malec, 1992). Subspecialization will not provide an antidote to diminishing professional resources. Factionalism is a frequent response within an organization or profession when the environment within which the organization functions suddenly changes (Balser, 1997). It is a response that must be met with increased leadership, communication among the splintering groups, and a re-focusing upon a unified mission.

In order for new models of education, practice, and theory in clinical psychology to develop and thrive, new leaders in the field will need to be actively fostered. The profession of psychology tends to recycle its leadership, with a small number of individuals repeatedly garnering the decision-making and policy-making positions (Cummings, 1996, p. 7). New leadership is not visibly fostered. This may be due to the lack of a career ladder in psychology (Sheridan, 1999). Political advancement within the bureaucracy of the profession is one of the few clearly outlined means of professional advancement, a situation that could contribute to a recycling of leaders and bottlenecking of new ideas. The profession's leadership must begin to actively promote and develop upcoming generations of innovators if new ideas are to be heard and tried. Without such encouragement of a larger number of new leaders, the pace of change within the profession will continue to stagnate. One advantage of developing hybrid professionals is that they can bring back to psychology successful ideas and professional strategies learned from other professional affiliations, possibly resulting in a very diversified and broad leadership.

The interdisciplinary and cross-training model of developing clinical psychology graduates capable of hybrid careers may be particularly well suited for terminal masters programs in psychology. The masters level clinical psychologist has been virtually disowned by most national and many state psychological associations (Kohout, 1993). Masters level psychology graduates have little to gain from exclusive adherence to, and identification with, traditional clinical roles within the profession. The masters level psychology graduate who is well prepared to pursue other career paths, however, will be able to benefit from advanced psychological knowledge and skills, free from the constraints of a profession in which he has little control. In fact, terminal masters level programs in psychology, because they have less regulated curriculum requirements, may be able to adapt their curricula more rapidly to facilitate interdisciplinary programs.

The robust achievements of the discipline of clinical psychology must be preserved. One way to do this is to allow the discipline to be applied in innovative ways and in new professional settings. By creating hybrid professionals with expanded career opportunities, graduate education in clinical psychology will achieve a new model of professional preparation that is consistent with the needs of the contemporary job market. It is the ultimate act of "giving psychology away" (Miller, 1969). By creating employable graduates, psychology will be able to retain its claim as a helping profession, since adaptation will have been made for the benefit of those students who entrust a significant portion of their time and resources to study the discipline. By educating students to address diverse problems, rather than adhere to a profession or "guild" (Kiesler, 1987, p. 47), professional clinical psychology will actually prevail, since education in psychology will have conquered the limits imposed by academic disciplines and professional borders in order to serve the broader community.

REFERENCES

American Association of Oriental Medicine (1998). *Educational options in oriental medicine*. Catasauqua, PA: Author.

American Physical Therapy Association (1990). A future in physical therapy. Alexandria, VA: Author.

Angell, M., & Kassirer, J. P. (1998). Alternative medicine: the risks of untested and unregulated remedies [editorial]. *New England Journal of Medicine, 339*(12), 839-841.

Arthur, W., & Benjamin, L.T. (1999). Psychology applied to business. In: A.M. Stec, D.A. Bernstein, et al (Eds.) *Psychology: Fields of Application*. Boston: Houghton Mifflin.

Barlow, W. (1990). *The Alexander technique*. New York: Knopf Publishers.

Basler, D. (1997). The impact of environmental factors on factionalism and schism in social movement organizations. *Social Forces, 76*(1), 199-228.

Belar, C. D. (1997). Clinical health psychology A specialty for the 21st century. *Health Psychology, 16*(5), 411-416.

Belar, C. D. (1995). Collaboration in capitated care: Challenges for psychology. *Professional Psychology: Research and Practice, 26*, 139-146.

Bent, R.J., Packard, R.E., & Goldberg, R.W. (1999). The American Board of Professional Psychology, 1947 to 1997: a historical perspective. *Professional Psychology: Research and Practice*, *30*(1), 65-73.

Berger, P., Hawley, N.P., & Pincus, J. (1992). Health and healing: Holistic practices and systems. In: The Boston Women's Health Book Collective (Eds.) *The New Our Bodies, Ourselves* (pp. 79-99). New York: Simon and Schuster.

Bersoff, D.N., Goodman-Delahunty, J., Grisso, J.T., Hans, V.P., Poythress, N.G., & Roesch, R.G. (1997). Training in law and psychology: Models from the Villanova conference. *American Psychologist*, *52*(12), 1301-1310.

Boucher, T.A., & Lenz, S.F. (1998). An organizational survey of physicians' attitudes about and practice of complentary and alternative medicine. *Alternative Therapies*, *4*(6), 59-65.

Boyer, E.L. (1990). Scholarship reconsidered: Priorities of the professoriate. New Jersey: The Carnegie Foundation for the Advancement of Teaching.

Brauzer, B., Lefley, H.P., & Steinbook, R. (1996). A module for training residents in public health systems and community resources. *Psychiatric Services, 47*(2), 192-194.

Bray, J.H. (1996). Psychologists as primary care providers. In: R.H. Rozensky and R.J. Resnick (Eds.) *To Your Health: Psychology Across the Lifespan* (pp. 89-100). Washington, D.C: American Psychological Association.

Brentar, J., & McNamara, J.R. (1991). The right to prescribe medication: Considerations for professional psychology. *Professional Psychology: Research and Practice*, *22*, 179-187.

Brink, S. (1999). New programs teach teamwork: Health care students gain new respect for colleagues in other disciplines. In: *U.S. News and World Report*: Best Graduate Schools (p. 45). Washington, D.C: U.S. News and World Report.

Broskowski, A., Marks, E., & Budman, S.H. (1981). *Linking health and mental health.* Beverly Hills: Sage Publications.

Cassileth, B.R. (1998). *The alternative medicine handbook.* New York: W.W. Norton.

"Changing name in a changing world...College of Health Professions" (1995, November 17). *Friday Evening Post Newsletter* of the University of Florida Health Sciences Center, p. 8.

Cook, S.W. (1970). Introduction. In: F.F. Korten, S.W. Cook, & J.I. Lacey (Eds.) *Psychology and the problems of society.* Washington, D.C.: American Psychological Association. (Pp. ix-xx).

Cooper, R.A.., Henderson, T., Dietrich, B.S. (1998). Roles of nonphysician clinicians as autonomous providers of patient care. *JAMA*, 280(9), 795-802.

Cooper, R.A., Laud, P., Dietrick, B.S. (1998). Current and projected work force of non-physician clinicians. *JAMA, 280*(9), 788-794.

Cordes, D.H., Rea, D.F., Rea, J.L., Vuturo, A.F., & Kligman, E.W. (1996). A survey of residency management training: general preventative medicine graduates. *American Journal of Preventive Medicine*, *12*(3), 172-176.

Cox, R.H. (1997, June). *Psychology in a new key*. Paper presented at the Missouri Psychological Association Convention, Lake Ozark, Missouri.

Cummings, N.A. (1996). The resocialization of behavioral healthcare practice. In: N.A. Cummings, M.S. Pallak, & J.L. Cummings (Eds.) *Surviving the Demise of Solo Practice: Mental Health Practitioners Prospering in the Era of Managed Care*. Madison: Psychosocial Press (pp. 3-10).

Cummings, N.A., & Duhl, L.J. (1987). The new delivery system. In: L.J. Duhl & N.A. Cummings (Eds.) *The Future of Mental Health Services: Coping With Crisis* (pp. 85-98). New York: Springer Publishing Co.

Davidson, R.H., & Oleszek, W.J. (1998). Congress and its members (6th Ed.). Washington, D.C: *Congressional Quarterly Press*.

DeLeon, P.H. (1988). Public policy and public service: Our professional duty. *American Psychologist, 43*(4), 309-315.

DeLeon, P.H. (1993). "On top of the hill": Indiana, national health and the DOD training program. *Register Report, 19*(3), 8.

DeLeon, P.H., Kjervik, D., Kraut, A.G., & VandenBos, G.R. (1985). Psychology and nursing: A natural alliance. *American Psychologist, 40*(11), 1153-1164.

DeLeon, P.H., & Pallak, M.S. (1982). Public health and psychology: An important, expanding interaction. *American Psychologist, 37*(8), 934-935.

Drucker, P.F. (1990). *Managing the nonprofit organization*. New York: Harper Collins.

Eisenberg, D.M., Kessler, R.C., Foster, C., Norlock, F.E., Calkins, D.R., & Delbanco, T.L. (1993). Unconventional medicine in the United States: Prevalence, costs, and patterns of use. *New England Journal of Medicine, 328*, 248-252.

Epstein, J., Sage, L., & Weding, D. (1995). A multimedia program to educate the public about mental and addictive disorders. *Behavior Research Methods, Instruments, and Computers, 27*(2), 289-292.

Erffmeyer, E.S., & Mendel, R.M. (1990). Masters level training in industrial/organizational psychology; A case study of perceived relevance of graduate training. *Professional Psychology: Research and Practice, 21*(5), 405-408.

Fairbanks, J., & Wiese, W.H. (1998). *The public health primer*. Thousand Oaks, California: Sage Publications.

Faust, D., Guilmette, T.J., Hart, K.J., Arkes, H.R., Fishburne, F.J., & Davey, L. (1988). Neuropsychologists' training, experience, and judgement accuracy. *Archives of Clinical Neuropsychology, 3*, 145-163.

Finley, N.J. (1995). Psychology in context: Making connections to other disciplines. *Teaching of Psychology, 22*(2), 105-108.

Fisher, H.S. (1997). *American salaries and wages survey* (4th Ed.) Detroit: Gale Publishing.

Folen, R.A., Kellar, M.A., James, L.C., Porter, R.I., & Peterson, D.R. (1998). Expanding the scope of clinical practice: The physical exam. *Professional Psychology: Research and Practice, 29*(2), 155-159.

Freudenheim, M. (1997, November 2). The future wears white: Nurses treading on doctors' turf. *.The New York Times*.

Gillette, N.P. (1998). A vision for our future. *American Journal of Occupational Therapy, 52*(5), 318-319.

Goleman, D., & Gurin, J. (1993). What is mind/body medicine? In: D. Goleman & J. Gurin (Eds.) *Mind/Body Medicine* (pp3-18). New York: Consumer Reports Books.

Goldberg, L.R. (1968). Simple models or simple processes? Some research on clinical judgements. *American Psychologist, 23*, 483-496.

Gorman, B. (1996). Behavioral practice: Notes and observations from a managed behavioral health executive. In: N.A. Cummings, M.S. Pallak, & J.L. Cummings (Eds.) *Surviving the Demise of Solo Practice: Mental Health Practitioners in the Era of Managed Care.* (Pp. 41-51). Madison: Psychosocial Press.

Griffith, J.R. (1995). *The well-managed health care organization* (3rd Ed.). Ann Arbor: AUPHA/Health Administration Press.

Haley, W.E., McDaniel, S.H., Bray, J.H., Frank, R.G., Heldring, M., Johnson, S.B., Lu, E.G., Reed, G.M., & Wiggins, J.G. (1998). Psychological practice in primary care settings: Practical tips for clinicians. *Professional Psychology: Research and Practice, 29*(3), 237-244.

Halpern, D.F., Smothergill, D.W., Allen, M., Baker, S., Baum, C., Best, D., Ferrari, J., Geisinger, K.F., Gilden, E.R., Hester, M., Keith-Spiegel, P., Kierniesky, N.C., McGovern, T.V., McKeachie, W.J., Prokasy, W.F., Szuchman, L.T., Vasta, R., & Weaver, K.A. (1998). Scholarship in psychology: A paradigm for the twenty-first century. *American Psychologist, 53*(12), 1292-1297.

Hayes, N. (1996). What makes a psychology graduate distinctive? *European Psychologist, 1*(2), 130-134.

Holland, D. (1998). The cost effective delivery of rehabilitation psychology services: the responsible utilization of paraprofessionals. *Rehabilitation Psychology, 43*(3), 232-245.

Horrigan, B. (1998). Eugene Taylor on spiritual healing and the American visionary tradition. *Alternative Therapies in Health and Medicine, 4*(6), 79-87.

Houston, B. (1996). Bury the liberal vs. professional arts debate. *Education, 117*(1), 12-16.

Johnson, J.A. (1998). Are Chinese medical qiqong standards too high for the west? *Qi: The Journal of Traditional Eastern Health and Fitness, 8*(4), 36-41.

Kaas, M.J., Dahl, D., Dehn, D., Frank, K. (1998). Barriers to prescriptive practice for psychiatric/mental health clinical nurse specialists. *Clinical Nurse Specialist, 12*(5), 200-204.

Keane, A., Richmond, T., & Kaiser, L. (1994). Critical care nurse practitioners: evolution of the advanced practice nursing role. *American Journal of Critical Care, 3*(3), 231-237.

Keita, G.P., & Jones, J.H. (1990). Reducing adverse reaction to stress in the workplace: Psychology's expanding role. *American Psycholgist, 45*(10), 1137-1141.

Kiesler, C.A. (1987). The guilds: An organizational analysis. In: L.J. Duhl & N.A. Cummings (Eds.) *The future of mental health services: Coping with crisis* (pp. 55-70). New York: Springer Publishing Co.

Klein, J. (1990) _*Interdisciplinarity: History, theory, and practice*. Detroit: Wayne State University Press.

Klusman, L.E. (1998). Military health care providers' view on prescribing privileges for psychologists. *Professional Psychology: Research and Practice, 29*(3), 223-229.

Kockelmans, J. (1979). *Interdisciplinarity and higher education*. University Park: Pennsylvania State University Press.

Kohout, J. (1993). *Employment opportunities in psychology: Current and future trends*. Paper presented at the American Psychological Association Annual Convention, Toronto, August.

Koop, C.E. (1983). Perspectives on future health care. *Health Psychology, 2*(3), 303-312.

Krauss, D.A., Ratner, J.R., & Sales, B.D. (1997). The antitrust, discrimination, and malpractice implications of specialization.

Lasker, R.D., & The Committee on Medicine and Public Health (1997). *Medicine and public health: The power of collaboration*. New York: The New York Academy of Medicine.

Last, J. M. (Ed.) (1995). *A dictionary of epidemiology* (3rd Ed.). New York: Oxford Press.

Lee, J.A., DeLeon, P.H., Wedding, D., & Nordal, K. (1994). Psychologists' role in influencing congress: The process and the players. *Professional Psychology: Research and Practice, 25*(1), 9-15.

Leli, D.A., & Filskov, S.B. (1981). Clinical-actuarial detection and description of brain impariment with the W-B Form I. *Journal of Clinical Psychology, 37*, 623-629.

Leli, D.A. & Filskov, S.B. (1984). Clinical detection of intellectual deterioration associated with brain damage. *Journal of Clinical Psychology, 40*, 1435-1441.

Lutz, R.A. (1996). The higher education system: Liberal arts and the business world. *Vital Speeches of the Day, 62*(21), 649-652.

Maddi, S. (1997). Strengths and weaknesses of organizational consulting from a clinical psychology background. *Consulting Psychology Journal: Practice and Research, 49*(3), 207-219.

Malec, J. (1992). Consumer protection in the expansion of neuropsychology. *Counseling Psychologist, 20*(4), 620-625.

Matarazzo, J. (1983). Education and training in health psychology: Boulder or bolder. *Health Psychology, 2*(1), 73-113.

Mattis, S. (1999). A message from the president. *The Diplomate, 18*(2), 2.

McCormick, S. (1994). Public administrators: Managers who know how to operate government. In: B.J. Morgan & J.M. Palmisano (Eds.) *Public Administration Career Directory*. Detroit: Gale Publishing (pp 1-4).

McCutcheon, L. (1995, July/August). Bach flower remedies: time to stop smelling the flowers? *Skeptical Inquirer, 55*, 33-35.

McCutcheon, L. (1994). Homeopathic remedies for psychological problems: Reconciling the claims with the evidence. *Journal of Applied Nutrition, 46*, 86-91.

Menkin, D. (1999). A simpler way. *Massage Magazine*, January/February, 106-111.

Miller, G.A. (1969). Psychology as a means of promoting human welfare. *American Psychologist, 24*, 1063-1075.

Moore, N.G. (1997). A review of reimbursement policies for alternative and complementary therapies. *Alternative Therapies in Health and Medicine, 3*(1), 26-29.

Moss, M.T. (1996). Preparing nurse managers for a managed care future. *Nursing Economics, 14*(2), 132-133.

Mourad, R.P. (1997). Postmodern interdisciplinarity. *The Review of Higher Education, 20* (2), 113-140.

Muscat, M. (1999). OAM elevated to center status. *Alternative Therapies in Health and Medicine, 5*(1), 24-25.

Neistadt, M.E., & Crepeau, E.B. (1998). *Occupational therapy* (9th Ed.). Philidelphia: Lippincott Publishers.

Ogloff, J.R.P. (1999). Graduate training in law and psychology at Simon Fraser University. *Professional Psychology: Research and Practice, 30*(1), 99-103.

Offerman, L.R, & Gowing, M.K. (1990). Organizations of the future: Changes and challenges. *American Psychologist, 45*(2), 95-108.

Pearson, L.J. (1999). Annual update of how each state stands on legislative issues affecting advanced nursing practice. *The Nurse Practitioner, 24*(1), 16-24.

Prieto, J.M., & Avila, A. (1994). Linking certified knowledge to labor markets. *Applied Psychology: An International Review, 43*(2), 113-130.

Quick, J.C. (1999). Occupational health psychology: The convergence of health and clinical psychology with public health and preventive medicine in an organizational context. *Professional Psychology: Research and Practice, 30*(2), 123-128.

Rabasca, L. (1998). Structuring departments for greater impact. *APA Monitor, 29*(12), p. 4.

Ranz, J., & Stueve, A. (1998). The role of psychiatrist as program medical director. *Psychiatric Services, 49*(9), 1203-1207.

Raymond, J. (1999). Beating those HMO blues. In: *Kaplan/Newsweek Careers 2000* (p. 26). New York: Newsweek, Inc.

Resnick, J.H. (1991). Finally, a definition of clinical psychology: A message from the president, Division 12. *The Clinical Psychologist, 44*, 3-11.

Resnick, R.J., DeLeon, P.H., & VandenBos, G.R. (1997). Evolution of professional issues in psychology: Training standards, legislative recognition, and boundaries of practice. In: J.R. Matthews and C.E. Walker (Eds.) *Basic Skills and Professional Issues in Clinical Psychology* (pp281-303).. Boston: Allyn and Bacon.

Ripley, B., & Franklin, G.A. (1991). *Congress, the bureaucracy, and public policy* (5th Ed.). Pacific Grove, CA: Brooks/Cole.

Roberts, M.C. (1998). Innovations in specialty training: The clinical child psychology program at the University of Kansas. *Professional_Psychology: Research and Practice, 29*(4), 394-397.

Robiner, W.N. (1991). Dialogue on a human resource agenda for psychology: A welcome and a response. *Professional Psychology: Research and Practice, 22*(6), 461-463).

Rosellini, L. (1999). Alternative goes mainstream: Medical school scramble to add courses on nontraditional therapies. In: *U.S. News and World Report*

Best Graduate Schools, 2000 Edition (pp. 47-48). Washington, D.C.: U.S. News and World Report.

Rosenthal, M.P., Diamond, J.J., Rabinowitz, H.K., Bauer, L.C., Jones, R.L., Kearl, G.W., Kelly, R.B., Sheets, K.J., Jaffe, A., Jonas, A.P., et al (1994). Influence of income, hours worked, and loan repayment on medical students' decision to pursue a primary care career. *JAMA, 271*(12), 914-917.

Rosenthal, M.P., Rabinowitz, H.K., Diamond, J.J., & Markham, F.W. (1996). Medical students' specialty choice and the need for primary care. Our future. *Primary Care, 23*(1), 155-167.

Sarafino, E.P. (1998). *Health psychology: Biopsychosocial interactions* (3rd. Ed.). New York: John Wiley & Sons, Inc.

Schaffer, C. (1998). The balanced budget act: Impact on providers. *Remington Report, 6*(3), 16-17.

Schippman, J.S., Hawthorne, S.L., & Schmitt, S.D. (1992). Work roles and training needs for the practice of industrial-organizational psychology at the masters and Ph.D. level. *Journal of Business and Psychology, 6*(3), 311-331.

Schneider, S.F. (1981). Where have all the students gone? Positions of psychologists trained in clinical/services programs. *American Psychologist, 36*(11), 1427-1449.

Segal-Isaacson, A. (1998). Forecasting the future of the CNS. In: L.J. Pearson (Ed.) 1998/1999 *Advanced practice nurse sourcebook.* Springhouse, PA: Springhouse Corp.

Sharp, N. (1999). 1999: The road ahead for NPs. *The Nurse Practitioner, 24*(2), 120-124.

Sheridan, E.P. (1999). Psychology's future in medical schools and academic health care centers. *American Psychologist, 54*(4), 267-271.

Siegel, M., & Doner, L. (1998). Marketing public health: Strategies to promote social change. Gaithersburg, Maryland: Aspen Publications.

Silbiger, S. (1993). *The ten day MBA.* New York: William Morrow Publishers.

Sleek, (1996, November). *APA Monitor.*

Stamm, B.H. (1998). Clinical applications of telehealth in mental health care. *Professional Psychology: Research and Practice, 29*(6), 536-542.

Starr, P. (1982). *The social transformation of American medicine.* New York: Basic Books.

Thomas, J. (1999). OAM to consider massage research initiative. *Massage Magazine*, January/February, 114-117.

Thomas, S.P. (1996). Nurse psychologists: A unique group within health psychology. *Journal of Clinical Psychology In Medical Settings, 3*(2), 93-101.

Troy, W.G., & Shueman, S.A. (1996). Program redesign for graduate training in professional psychology: The road to accountability in a changing professional world. In: N.A. Cummings, M.S. Pallak, & J.L. Cummings (Eds.) *Surviving the Demise of Solo Practice: Mental Health Practitioners in the Era of Managed Care* (pp 55-80).

United States Department of Labor, Bureau of Labor Statistics (1998). *Occupational outlook handbook*, 1998-99 Edition. Washington, D.C: Superintendent of Documents, U.S. Government Printing Office.

VandenBos, G.R., & DeLeon, P.H. (1998). Clinical geropsychology and U.S. federal policy. In: M. Hersen & V.B. Van Hasselt (Eds.) *Handbook of Clinical Geropsychology.* New York: Plenum Press. (Pp 19-28).

Wagman, M. (1998). *Cognitive science and the mind-body problem: From philosophy to psychology to artificial intelligence to imaging of the brain.* Westport, CT: Praeger Publishers.

Walley, P.B. (1995). Lucky dogs. *Professional Psychology: Research and Practice, 26*(5), 459-462.

Weatley, B., DeJong, G., & Sutton, J.P. (1997). Managed care and the transformation of the medical rehabilitation industry. *Health Care Management Review, 22*(3), 25-39.

Wiens, A.N. (1993). Postdoctoral education-training for specialty practice. *American Psychologist, 48*(4), 415-422.

Wiggins, J.G. (1994). Would you want your child to be a psychologist? *American Psychologist, 49*(6), 485-492.

Williams, S., & Kohout, J.C. (1999). Psychologists in medical schools in 1997: research brief. *American Psychologist, 54*(4), 272-276.

Winett, R.A., King, A.C., & Altman, D.C. (1989). *Health psychology and public health: an integrative approach.* New York: Pergamon Press.

World Bank (1993). *World development report 1993, investing in health.* New York: Oxford University Press.

Wu, W.H., Bandilla, E., Ciccone, D.S., Yang, J., Cheng, S.C., Camer, N., Wu, Y., & Shen, R. (1999). Effects of qiqong on late-stage complex regional pain syndrome. *Alternative Therapies in Health and Medicine, 5*(1), 45-54.

Yung, B.R., Hammond, W.R., Sampson, M., & Warfield, J. (1998). Linking psychology and public health: A predoctoral clinical training program in youth violence prevention. *Professional Psychology: Research and Practice, 29*(4), 398-401.

Zilberg, N., & Carmody, T.F. (1995). New directions for the education of clinical psychologists: The primary care setting, the VA's PRIME program, and the in-depth generalist model. *Journal of Clinical Psychology in Medical Settings, 2*(1), 109-127.

Book Review: *Guiding Your Child through Grief*

Ross Robak
Pace University

As school psychologists, we're occasionally confronted with children's reactions to a death in the family. Sometimes these losses are unrecognized. Our broad American culture disenfranchises grief responses of many kinds. One commonly unrecognized grief reaction is following the death of a family member such as a grandparent, or an aunt/uncle. At other times, as with the death of a parent or sibling, the loss is more clearly impacting. But even in such cases, many of us find it very difficult to know how to respond, what to suggest, or what to do to help the child and the family in grief. This is particularly significant in light of the well-documented fact that most young adults report retrospectively that their first significant loss was a death in the family.

Books in the death and grieving field tend to fall into one of two categories: They are either written to a professional audience and are often removed from everyday experiences or they are first person accounts of survivors. In the first case, these books are often theoretical and abstract. In the second, they are idiosyncratic and leave one wondering whether the writing wasn't simply the author's catharsis. Mary Ann and James Emswiler have combined these two genres with clear down-to-earth explanations of theory and numerous case studies, including their own personal experiences.

Guiding Your Child through Grief (Bantam, ISBN 0553380257) is a very useful book. It is written mainly for the parents of children who have lost someone to death. But, the book is also helpful to teachers, school psychologists, and counselors working with children. For example, the authors have managed (in chapter 2) to highlight the important literature on cognitive development and children's grieving. This includes:

Before age two, lacking the requisite language skills, the child may not conceptualize a loss. However, he or she experiences a vague sense of "gone-ness" based on his or her earliest recognition of parents by their smells, by the sounds of their voices, by the rhythms of interactions with each parent.

Between ages two and six, preschoolers' thoughts about a death tend to be marked by magical thinking (believing they have powers to have caused the death or to reverse it through wishing or praying). In addition,

---

Author info: Correspondence should be sent to: Dr. Rostyslaw W. Robak, Psych. Dept., Pace University, 861 Bedford Rd., Pleasantville, NY 10570.

North American Journal of Psychology, 2001, Vol. 3, No. 1, 171-172.
© NAJP

they usually struggle with three concepts: non-functionality of the body, finality (irreversibility) of the death, and universality (everybody dies).

Six- to nine-year-olds tend to have a better understanding of death, but it is still far from complete. They often understand that death is final but they may not recognize its universality (that they will die someday, too). They are often fascinated by the biological details of death, but they may, for example, think that death is contagious.

Between ten and thirteen, most children understand what "dead" means. However, during this period they begin to struggle with *why* death happens.

Adolescents can be said to think about death, in many respects, the way that adults do. What adults need to keep in mind is that adolescents struggle with some of the ultimate questions: "Why must *I* die?" "Why do bad things happen?" "What is the meaning of life and of death?"

Another set of questions for the school psychology professional is related to the nature of the grieving process itself. Parents and teachers often intuitively recognize that grief is a confusing experience. They are likely to have experienced a death in their own lives, perhaps even in their own childhood. The Emswilers have outlined five "challenges" that grieving children face. Adults can guide themselves by these challenges as they help a child through his or her grief. These challenges are:

> *Feeling safe in the world again.* For parents and teachers, this means: increasing predictability in the child's life; keeping one's own anxieties separate from the child's and talking with him or her about their health (and safety) concerns; helping them experience a sense of mastery in whatever areas they can; and being present (attentive and available).
>
> *Understanding the death.* Parents and others can: give accurate information in a concrete manner appropriate to the child's developmental level; read books together; and not trying to wipe out the loss.
>
> *Mourning the death.* This means, essentially, talking about feelings and facilitating other expressions of emotions (e.g., drawing, keeping a diary).
>
> *Staying connected with the person who died.* The often-cited Child Bereavement Study reported that 80% of bereaved children say that they feel as if the deceased person is watching them. It seems that children need to maintain their connections with the deceased. Keeping pictures, continuing projects that had been done with the dead person, and writing about them can serve to help children maintain their connections.
>
> *Resuming childhood again.* It's important to not let the grieving child lose his or her childhood. While children are facing all of the previous challenges, they are also children and thus need to play, have friends, and be allowed to fool around.

Guiding Your Child through Grief is filled with such insights. It is an exceptionally good bibliotherapy resource. Any adult involved in the life of a grieving child is bound to take many lessons from it.

Validity Comparison of the General Ability Measure for Adults with the Wonderlic Personnel Test[1]

J. Patrick Leverett
T. Darin Matthews
Kerry S. Lassiter
Nancy L. Bell
The Citadel

Future trends with regard to workplace demographics and the predicted multicultural expansion of the labor force are briefly reviewed, with a focus on the importance of cultural fairness in utilizing assessment in career counseling. The present study investigated the validity of the *General Ability Measure for Adults* (GAMA) and the *Wonderlic Personnel Test* (WPT) for a sample of 112 college students. Results indicated that GAMA IQ scores correlated significantly with WPT scores ($r = .61$, $p < 0.001$). Results suggest that the GAMA may provide utility in the assessment of individuals within career counseling settings who have limited command of English, have language difficulties (e.g., language-based learning disorders), or who possess handicaps which may limit their ability to effectively process linguistic stimuli (i.e., deafness). Implications for clinical practice, as well as future research, are discussed.

A number of issues relevant to the use of psychological assessment in career counseling should be considered by psychologists engaged in such activities. Intelligence exists as just one of the many issues pertinent to occupational choice, and admittedly might not ultimately prove as important as motivation, discipline, or other personality variables (i.e., ability to work and communicate with others, attitudes toward work, etc.). However, intellectual factors are generally assumed to be important in making broad decisions and choices within the process of career counseling. As such, assessment of the intellectual ability of clients (typically used to predict a client's potential for educational or occupational level) is an important aspect of career counseling (Lindeman & Matarazzo, 1990).

With an expanding multicultural population comprised of increasing numbers of persons with varying levels of English language proficiency,

---

Author info: Correspondence should be sent to: J. Patrick Leverett, Department of Psychology, The Citadel, 171 Moultrie St., Charleston SC 29409

North American Journal of Psychology, 2001, Vol. 3, No. 1, 173-182.
© NAJP

it is imperative that users of psychological tests consider the potential effects of language deficits in the testing process. In addition, the use of instruments that rely heavily on verbal content with persons who have language-based learning disorders may adversely impact assessment results. This study will investigate the validity of a relatively culture-fair instrument that makes use of minimal linguistic stimuli, specifically the *General Abilities Measure for Adults* (GAMA; Naglieri & Bardos, 1997). The GAMA will be compared to an instrument widely used in career counseling and occupationally-related decision making, the *Wonderlic Personnel Test* (WPT; Wonderlic Personnel Test, 1992).

Significant changes are expected to occur in the workplace over the next fifty years as a result of changing United States population demographics (Futurework, 1999). Predicted trends suggest that the North American workforce will continue to expand and to grow more diverse (Futurework, 1999; Latham & Sue-Chan, 1997), leading to a number of increasingly important areas of concern for psychologists and other counseling professionals who utilize psychological testing in career-related or occupational counseling. In 1995, White Americans constituted 73.6% of the US population. By 2050, White Americans are projected to comprise only 52.8% of America's populace, with Black and Native American segments of the U.S. population projected to show modest proportional population increases over this same period of time (12.0% to 13.6%, 0.7% to 0.9%, respectively). The most dramatic projected proportional increases in population levels involve groups specified as "Hispanic origin" and "Asian and Pacific Islander." Persons of Hispanic origin comprised 10.2% of America's population in 1995. By 2050, Hispanic Americans are projected to constitute 24.5% of the U.S. population. The Asian and Pacific Islander group is expected to grow in that same time period from a group comprising 3.3% of Americans to 8.2% of the U.S. population. Nearly two-thirds of the persons responsible for this population growth will be immigrants (Futurework, 1999).

Latham and Sue-Chan (1997) suggest that a number of aspects, including validity, reliability and freedom from bias, will be emphasized in selecting employees for the 21st century North American workplace. Among the most important of these issues are the potential discriminatory effects of language-based instrumentation with persons who may be entirely suitable for certain occupational settings, but who score poorly on tests due to their relative inefficiency of language usage. The issue of potential discrimination in assessing those for whom English is a non-preferred language must be addressed and adequately resolved. This is particularly important when such applicants are considering occupations for which English language proficiency is not an essential aspect of job performance.

In addition to issues related to the acquisition of English language, other factors also have significant impact with regard to the appropriate use of psychological assessment in career counseling. Individuals who have language-based learning disorders may perform less efficiently on tests which rely heavily on linguistic stimuli. Similarly, individuals who are deaf or who have significant hearing impairment may have limited command of written language due to their use of an alternative form of language (i.e., American Sign Language). These individuals may be entirely well-suited for a number of jobs, but may demonstrate poor performance on intelligence tests that are comprised of verbal content and that rely heavily on language skills.

These points emphasize a number of issues pertinent to psychological assessment of clients involved in career counseling. Standards for appropriate use of psychological tests stress that psychologists consider the cultural backgrounds and prior experiences of ethnicity, age, gender and cultural groups (Standards for Educational and Psychological Testing, 1985). Indeed, linguistic and experiential differences are essential elements of possible bias when traditionally American instruments are used to assess culturally different individuals.

In an attempt to address this issue, several psychological instruments have been translated into alternative-language formats. Simple translation is not in itself sufficient, since virtually all languages have differing expression of concepts and phrases, unfamiliar idiomatic usage, etc. Even when reliable and valid alternative language versions of instruments exist, it is a dilemma to understand exactly what mechanisms may predicate the differential observed performance of majority versus minority cultures (especially those for whom English is a preferred language and those for whom it is not). For example, Helms (1992) described the inherent "Eurocentricism" in cognitive ability testing, pointing out how culture may influence test responses in a variety of ways. Perez-Arce (1999) described the "cultural boundness" of cognition, pointing out that language and socioeconomic factors are especially influential in the way that an individual responds on cognitive ability testing. Making matters even more complex, the equivalence of psychometric properties including item content, reliability, validity, and level of difficulty cannot be assumed when instruments are translated from one language to another (Standards for Educational and Psychological Testing, 1985).

The Wonderlic Personnel Test (WPT) is perhaps the most widely used instrument to assess intellectual ability in career and occupationally-related counseling and assessment settings. In addition to the original

English language version, the WPT has been translated into 12 different languages. With the exception of the Spanish language version, the other alternative language versions of the WPT are currently considered developmental due to insufficient reliability and validation research (Wonderlic, 1992). The WPT is purported to be a good measure of general intelligence, or "g" (Wonderlic, 1992) and has been shown in numerous studies to be correlated with the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1955) and Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981) (Dodrill, 1981, 1983; Dodrill & Warner, 1988; Hawkins, Faraone, Pepple, Seidman & Tsuang, 1990).

In addition to issues related to the appropriate use of psychological tests in career counseling, this study was conducted to contribute to the growing amount of validation research concerning the GAMA. Naglieri and Bardos (1997) have reported results from a number of studies examining the validity of the GAMA. These include reported significant and strong positive relationships between the GAMA IQ score and the

WAIS-R (Wechsler, 1981), the *Kaufman Brief Intelligence Test* (K-BIT; Kaufman & Kaufman, 1990), and the WPT (Wonderlic, 1992). Review of these studies suggests that the GAMA provides a good estimate of general intellectual ability (Naglieri & Bardos, 1997). In examining the relationship between the GAMA, the WPT, and the *Shipley Institute of Living Scale* (Shipley, 1967), Naglieri and Bardos (1997) administered these instruments to a sample of 80 individuals ranging in age from 25 to 45 ($M = 34.2$, $SD = 6.2$). They found that the GAMA IQ score correlated positively with the WPT ($r = .70$, $p < .001$); significant positive correlations were also found between the GAMA and the Shipley Vocabulary (.56), Shipley Abstraction (.73), and Shipley Total Scores (.72).

A small number of additional validation studies have been conducted investigating the validity of the GAMA. Lassiter, Leverett, and Safa (2000) found that the GAMA measures nonverbal intelligence to a greater degree than it does general intelligence. They reported that the statistically significant positive correlation between the GAMA and WAIS-R Performance IQ was itself significantly greater than the significant positive correlations observed between the GAMA and WAIS-R Verbal and Full Scale IQ scores. Their findings suggest that while the GAMA does measure verbal aspects of intelligence and can be viewed as a measure of general intelligence, it is more clearly associated with nonverbal, perceptual-organizational skills.

Clearly, additional research is needed to provide evidence of the GAMA's validity, both in general terms as well as its validity for specific uses (i.e., career counseling). In this study, the concurrent validity of the GAMA was investigated, using the WPT as the criterion measure.

METHOD

Participants and Procedure

Study participants were 112 adult college students attending a small southeastern college. Fifty-five women and 57 men participated in the study (77.7% Caucasian, 12.5% African-American, 4.5% Hispanic, 3.6% Asian, .9% Pacific Islander, and .9% Other). Participants' ages ranged from 18 to 55 years (*M* = 26.9, *SD* = 8.60). Participants were administered the GAMA and WPT according to the procedures described in the respective test manuals. The WPT score was converted to a metric with a mean of 100 and a standard deviation of 15 by subtracting the normative mean of 21.06 and dividing the normative standard deviation of 7.12 from each individual's score to obtain the standard score. (Wonderlic, 1992). The GAMA and WPT were part of a larger battery of intelligence measures administered to participants.

Instruments

The *General Abilities Measure for Adults* (GAMA; Naglieri & Bardos, 1997) is a paper and pencil intelligence test which measures reasoning ability through the use of abstract geometric designs. The GAMA takes 25 minutes to complete and is comprised of four subscales: Matching, Analogies, Sequences, and Construction. Examinees must choose the best response from among six alternatives. For example, on the Matching subscale, the examinee must choose from the alternative responses the abstract geometric design which best matches the stimulus item. The response chosen must match the stimulus item according to color, shape, and configuration.

The Analogies sub-scale involves presentation of an abstract design pair which is followed by another design pair in which the last design is missing. Examinees must study one stimulus design pair and choose the missing component which will complete the second stimulus design pair from a number of response alternatives. The examinee must recognize the relationship between the design in the first design pair and then select the design that best perpetuates this conceptual relationship in completing the second design pair.

The Sequences subscale consists of geometric designs presented in sequential frames. Design shapes, colors, and locations within each frame change in a logical sequence from left to right, with one frame within the sequence left blank. The examinee must recognize the pattern and choose the design which best completes the pattern sequence.

The Construction subscale requires that the examinee mentally configure several disconnected shapes of differing colors in order to produce a design. The task requires that the examinee analyze and mentally consolidate the spatial aspects of the colored shapes in order to construct a design to match one of the response alternatives.

The *Wonderlic Personnel Test* (WPT; Wonderlic Personnel Test, 1992) is a paper and pencil short form test of general cognitive ability. The WPT consists of 50 questions which incorporate a variety of types of problems. WPT questions include word comparisons, disarranged sentences, sentence parallelism, following directions, number comparisons, number series, geometric figure analysis, and mathematics and logical story problems. Items begin at a modest difficulty level, and gradually increase in difficulty. The WPT can be administered in groups, with examinees given 12 minutes to see how many items they can correctly answer.

RESULTS

Means, standard deviations, and Pearson product-moment correlations for the GAMA IQ score and subscales (Matching, Analogies, Sequences, and Construction) and the WPT are presented in Table 1.

TABLE 1 Means, Standard Deviations, and Correlations for the GAMA and Wonderlic Personnel Test (WPT)

Measure	Mean	*SD*	WPT r	r[d]
WPT[a]	106.19	11.05		
GAMA IQ[b]	105.35	10.77	.61**	.73
GAMA Matching[c]	11.48	2.59	.31*	.35
GAMA Analogies[c]	11.13	1.96	.39**	.54
GAMA Sequences[c]	10.99	2.38	.54**	.62
GAMA Construction[c]	9.91	3.05	.50**	.50

Note. For all variables, N = 112. [a]*Wonderlic Personnel Test Score* is converted to a metric with a mean of 100 and a standard deviation of 15 by subtracting the normative mean of 21.06 and dividing the normative standard deviation of 7.12 from each individual's score to obtain the standard score. (Normative information obtained from the *Wonderlic Personnel Test and Scholastic Level Exam User's Manual*, 1992).

[b]GAMA IQ Score is expressed with a mean of 100 and a standard deviation of 15.

[c]GAMA subscale scores are expressed with a mean of 10 and a standard deviation of 3.

[d]Correlation corrected for range restriction.

*$p < 0.01$

**$p < 0.001$

The mean WPT Total Score and GAMA IQ score were slightly higher than 100; this is not unexpected given that participants were all college students. Pearson correlation coefficients presented in Table 1 show that the GAMA IQ score was significantly positively correlated with the WPT total score. Furthermore, the GAMA Matching, Analogies, Sequences, and Construction subscales also were significantly positively correlated with the WPT. Correlations were corrected for restriction of range (Guilford & Fruchter, 1978; see Table 1 for corrected correlations), yet demonstrated this same pattern of performance.

To further investigate the validity of the GAMA, we examined the degree to which GAMA IQ scores predict individual performance on the WPT. In most clinical settings, it is this degree of predictive capacity that is essential. We calculated the percent of absolute difference scores between the GAMA and WPT. Absolute difference scores presented in Table 2 demonstrate that 41% of participant's GAMA scores were within 5 points of their WPT scores and 75% were within 10 points. Ninety percent of GAMA scores were within 15 points of participant's WPT scores. It is noteworthy that 25% of participant's GAMA scores deviated by more than 10 points from their WPT score, suggesting that GAMA IQ scores may provide conflicting information when evaluated on a case-by-case basis against the WPT IQ scores.

TABLE 2 Cumulative Percentage of Absolute Value Difference Scores Between GAMA and WPT

Absolute difference scores	GAMA %	Absolute difference scores	GAMA %
0	4	11	79
1	9	12	83
2	22	13	85
3	26	14	88
4	34	15	90
5	41	16	91
6	46	18	95
7	52	23	96
8	59	25	97
9	71	26	98
10	75	27	99
(continued in two columns on the right)		28	100

Note. GAMA = *General Abilities Measure for Adults*; WPT = *Wonderlic Personnel Test.* Horizontal lines reflect 5, 10, and 15 point discrepancies, respectively.

DISCUSSION

Our findings suggest that the GAMA and WPT are significantly positively correlated, and our conclusion is that the GAMA provides a good brief measure of general intellectual ability. The GAMA IQ Score and each of the GAMA subscales were significantly positively correlated with the WPT. It was noted that GAMA Sequences, Construction, and Analogies subscales were more strongly positively correlated with the WPT than was the Matching subscale. This may suggest that while the

GAMA and WPT may assess some common abilities (general intelligence; spatial and sequential reasoning; analysis, synthesis and rotation of spatial designs, and the analysis of specific figural-spatial details), the WPT may not assess abilities underlying the recognition of similar conceptual figural relationships as strongly as the GAMA.

As our nation becomes more culturally pluralistic, and as we develop greater sensitivity regarding the rights of those who may possess linguistically-handicapping conditions to be fairly assessed and considered for employment, the use of non-discriminatory assessment procedures becomes increasingly important. Although the WPT is perhaps the most widely used brief intelligence scale currently utilized for career counseling and occupationally-related decision making, a high degree of language-based items are contained in the format of the test. Indeed, all WPT items require reading ability in order for the examinee to understand the instructions for completing the item, even the primarily "figural" items (i.e., those items requiring cognitive analysis of geometric designs, etc.) In addition, the majority of the WPT figural items (items 7, 38, 42, 49) occur late in the format of this 50-item timed test. As such, examinees with inefficient or impaired reading ability for any of a number of reasons may not reach many of these items within the WPT 12 minute testing period. Thus, the GAMA may provide utility in the brief assessment of general intellectual ability for individuals for whom assessment with instruments utilizing primarily language-based stimulus items would be inappropriate or undesirable.

We believe that the GAMA, given in conjunction with the WPT, can provide a valuable source of additional understanding with regard to a potential employee's intellectual ability. With regard to clients who have reduced language proficiency for any of a number of reasons already discussed, information provided by the GAMA may provide career counselors an enhanced perspective for viewing the intellectual ability of such applicants. This conclusion is supported by recent research by Lassiter, Hutchinson, Bell and Matthews (1999). These researchers examined the concurrent validity of the GAMA and the Wechsler Adult Intelligence Scale – Third Edition (WAIS-III; Wechsler 1997) and found significant correlations between the GAMA and WAIS-III Verbal IQ and Full Scale IQ scores. The strongest correlation in their study, similar to the findings of Lassiter, Leverett, and Safa (2000), was between the GAMA and the Perceptual Organization Index score. A contrasting viewpoint was also noted by Matthews, Lassiter, and Habedank (1998), who explored the concurrent validity of the GAMA by comparing it to the Shipley. They suggested that, in contrast to Naglieri and Bardo's (1997) conclusion that the GAMA is a measure of general intellectual ability, that the GAMA may best assess verbal, rather than abstraction, skills. Perhaps more research within actual career counseling populations can help to clarify these issues.

Certain limitations of this study must be noted. This study was based upon the performance of a college student sample. While the mean GAMA IQ score and the WPT Total score of the sample were only slightly higher than would be expected in the normal population, future studies might explore the relationship of the GAMA and WPT using a sample of individuals with less education. It must also be emphasized that suggestions regarding the potentially beneficial use of the GAMA as a measure of general intelligence with persons who have reduced language proficiency are speculative; future research exploring the utility of the GAMA within such populations should be conducted. It should also be noted that, ironically, given our discussion of multicultural issues, our sample was largely Caucasian (i.e., 77.7%). Future research should be conducted using a more racially heterogeneous sample.

REFERENCES

Dodrill, C. B. (1981). An economical method for the evaluation of general intelligence in adults. *Journal of Consulting and Clinical Psychology, 49*(5), 668-673.

Dodrill, C. B. (1983). Long term reliability of the Wonderlic Personnel Test. *Journal of Consulting and Clinical Psychology, 51*(2), 316-317.

Dodrill, C. B., & Warner, M. H. (1988). Further studies of the *Wonderlic Personnel Test* as a brief measure of intelligence. *Journal of Consulting and Clinical Psychology, 56*, 145-147.

Futurework: Trends and challenges (1999). [On-line]. Available: http://www.ilr.cornell.edu/library/e_archive/futurework/

Guilford, J. P., & Fruchter, B. (1978). *Fundamental statistics in psychology and education.* (6th ed.) New York: McGraw-Hill.

Hawkins, K. A., Faraone, S. V., Pepple, J. R., Seidman, L. J., & Tsuang, M. T. (1990). WAIS-R validation of the *Wonderlic Personnel Test* as a brief intelligence measure in a psychiatric sample. *Psychological Assessment, 2*(2), 198-201.

Helms, J. (1992). Why is there no study of cultural equivalence in standardized cognitive ability testing? *American Psychologist, 47*(9), 1083-1101.

Kaufman, A. S., & Kaufman, N. L. (1990). *K-BIT (Kaufman Brief Intelligence Test) manual.* Circle Pines, MN: American Guidance Service, Inc.

Lassiter, K. S., Hutchinson, M. B., Bell, N., & Matthews, T. D. (1999*). Validity of the General Ability Measure for Adults and the WAIS-III.* Manuscript submitted for publication.

Lassiter, K. S., Leverett, J. P., & Safa, T. A. (2000). A comparison between the General Ability Measure for Adults and the WAIS-R for learning problem adults. *Assessment, 7*(1), 63-72.

Latham, G. P., & Sue-Chan, C. (1997). Selecting employees in the 21st century: Predicting the contribution of I-O psychology. *Canadian Psychology, 39*(1-2), 14-22.

Lindemann, J. E., & Matarazzo, J. D. (1990). Assessment of adult intelligence. In G. Goldstein, & M. Hersen (Eds.), *Handbook of psychological assessment* (pp. 79-101). New York: Pergamon Press.

Matthews, T. D., Lassiter, K. S., & Habedank, H. (1998). *Concurrent validity of the General Ability Measure for Adults.* Manuscript submitted for publication.

Naglieri, J. A., & Bardos, A. N. (1997). *General Ability Measure for Adults.* Minneapolis, MN: National Computer Systems.

Perez-Arce, P. (1999). The influence of culture on cognition. *Archives of Clinical Neuropsychology, 14*(7), 581-592.

Standards for educational and psychological testing. (1985). Washington, D.C.: American Psychological Association.

Wechsler, D. (1955). *WAIS manual.* New York: The Psychological Association.

Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale - Revised.* New York: The Psychological Corporation.

Wechsler, D. (1997). *Manual for the Wechsler Adult Intelligence Scale (3rd ed.).* San Antonio, TX: The Psychological Corporation.

Wonderlic Personnel Test. (1992). *Wonderlic Personnel Test: User's manual for the WPT and SLE.* Liberty. IL: Wonderlic Personnel Test, Inc.

Shipley, W. C. (1967). *Shipley Institute of Living Scale.* Los Angeles, CA: Western Psychological Services.

---

[1] This research was supported by a Citadel Development Foundation Grant.

How to Submit Papers to *NA*JP

The submission of good papers is very important to any psychology journal. We hope that you will support our views about what a psychology journal should be like by submitting papers now. We will try to mail a decision on your paper within 5 weeks of receipt by us. We will publish your paper in an average of 5 months of receipt of a final accepted version.

Submit 1 paper copy of each manuscript prepared according to APA guidelines. Also include a diskette with your manuscript or (preferably) e-mail the file to NAJP@mpinet.net. Include 2 large stamped envelopes so we can send your manuscript out for review. We encourage you to have your paper read by peers before submission.

If the final paper is accepted, please send a check or purchase order for the amount specified with your final draft (billed at $25 dollars per typeset page in units of two pages).

Submission of a manuscript to the Journal is a contract to publish with the Journal if the paper is accepted. Because of our rapid editorial feedback, simultaneous submission to other journals is not allowed. We will offer brief comments on the suitability of a topic for publication in *NA*JP before a submission.

Send all correspondence to: *North American* Journal of Psychology, 240 Harbor Dr., Winter Garden, FL 34787, phone 407-877-8364.

SUBSCRIPTIONS

All subscriptions are on a calendar year basis only. If you order late in the year, the earlier issues will be sent to you. Institutional rate, $65. Individual rates, paid by check, $30. Add $7 for Canadian subscriptions. Add $8 for subscriptions outside North America. Phone orders accepted with Visa, MasterCard or American Express. Call 407-877-8364. The *North American* Journal of Psychology is now published semi-annually in June & December.

Visa ☐ AMEX ☐ MC ☐

Card # Exp. Date

Name

Address

City State Zip

A Summary of Our Philosophy

The ***North American*** **Journal of Psychology** publishes scientific papers that are of general interest to psychologists and other social scientists. Papers may encompass a wide variety of topics, methods, and strategies. It is our belief that many good scientific papers are rejected for invalid reasons. Some journals reject up to 90% of their submissions because they don't have space. Furthermore, studies of the review process strongly suggest that not all of the lucky 10% are better papers than those that get rejected. In our view many of those rejected papers are meritorious with need for only minor revisions. Perhaps our philosophy can be summed up this way: we are looking for ways to strengthen good papers rather than searching for weak excuses to reject them.

North American
Journal of Psychology

BEST COPY AVAILABLE

Vol. 3, No. 2 2001

Advisory Editorial Board

ADRIAN FURNHAM
Psychology
University College London

DOUGLAS JACKSON
Psychology
University of Western Ontario

JAMES M. JONES
Psychology
University of Delaware

HEIDI KELLER
Psychology
University of Osnabrueck

MANUEL LONDON
Psychology
SUNY, Stony Brook

KAREN LYNESS
Management
Baruch College, CUNY

ROBERT PASNAK
Psychology
George Mason University

RICHARD RYAN
Psychology
University of Rochester

PHILIP ZIMBARDO
Psychology
Stanford University

Editor
LYNN E. McCUTCHEON

How to Subscribe to The *North American* Journal of Psychology

All subscriptions are on a calendar year basis only. If you order late in the year, the earlier issue will be sent to you. Institutional rate: $65. Individual rate: $30. Add $7 for Canadian subscriptions. Add $8 for subscriptions outside North America. *North American* Journal of Psychology is published in March, June and December.

Now visit our website at http://najp.8m.com

North American Journal of Psychology

June/July 2001 ISSN 1527-7143 Vol. 3, No.2

Males' Impressions of Female Fans and Nonfans: There Really is "Something About Mary" 183
Daniel L.Wann, Jennifer Schinner & Brian L. Keenan

Parental Disciplinary History and Current Level of Empathy and Moral Reasoning Among Late Adolescents 193
N. L. Lopez, J. L. Bonenberger, & H. G. Schneider

An Interview with Russell Eisenman 205
Russell Eisenman & Michael F. Shaughnessy

Differential Processes of Emotion Space Over Time 217
Takuma Takehara & Naoto Suzuki

Healthy Perfectionism and Positive Expectations About Counseling 229
J. M. Oliver, Brett A. Hart, Michael J. Ross, & Barry M. Katz

Attitudes Toward Perpetration of Sexual Abuse Against Children: A Study of Legal & Social Service Oriented Students 243
Scott A. Hubbartt & Sangeeta Singg

Self-protective Pessimism: Optimistic Bias in Reverse 253
John Chapin

Lost Letter Technique: Returned Responses to Battered and Abused Women, Men, and Lesbians 263
F. Stephen Bridges,C. B. Williamson, P. C. Thompson, & M. A. Windsor

Filicide-suicide: In Search of Meaning 277
P. L. Collins, Michael F. Shaughnessy, Loretta Bradley, & K. Brown

Book Review: "Stop Smoking & Chewing Tobacco for Life Changes 293
John Karkut

The Role of Religion and Responses to the Y2K Macro-crisis 295
Larry Powell, Mark Hickson,III, William R. Self, & Jean Bodon

Stability of Self-esteem as Moderator of the Relationship Between Level of Self-esteem and Depression 303
A. F. de Man, B. I. Becerril Gutierrez, & N. Sterk

Legal Attitudes Scale: A Factor Analytic Study 309
James V. Couch & Donna L. Sundre

Life after College: Psychology Students' Perceptions of Salary, Business Hiring Criteria, and Graduate Admission Criteria 321
Deborah S. Brühl

The Effect of Politeness and Grammar on User Perceptions of Electronic Mail 331
Sheri L. Jessmer & David Anderson

Editor's Comments

This is the second issue for 2001, our third volume. NAJP was issued twice a year for the first two years, but this year we are doing three - this is the summer issue, which you are probably getting in July, to be followed by one in the late Fall.

Thanks again to all the reviewers, all of whom did a marvelous job of critiquing papers promptly and with great insight. Without our reviewers, this journal would simply not have the quality it has today.

This issue features an article by Daniel Wann, author of one of the leading textbooks in sport psychology, and an interview with Russell Eisenman, in which Russ discusses his research and his views on topics ranging from antisocial personality to the harm done to people when "political correctness" gets out of hand.

Recently, one of the Ivy League universities subscribed to NAJP. We would like your institution to follow suit. Please point out to your librarian that NAJP is relatively inexpensive and is designed to appeal to the majority of psychologists. If you can get your library to subscribe, we will offer you a free subscription.

Lynn E. McCutcheon, editor

Males' Impressions of Female Fans and Nonfans: There Really is "Something about Mary"

Daniel L. Wann, Jennifer Schinner, Brian L. Keenan
Murray State University

This experiment examined male sport fans' and nonfans' impressions of female fans. Participants with a high or low interest in sport fandom read a description of a target female presented as having a high or low interest in sport and a high or low interest in a local team. Participants then reported their impressions of the target. As expected, males with a strong interest in sport gave particularly positive ratings to the target described as possessing a similarly high interest. Participants with little interest in sport gave particularly high ratings to the target described as possessing both a high interest in sport and a high interest in a local team. These participants reported less favorable impressions when the target was described as having a high interest in sport but little interest in a local team.

Historically, sport fandom has been a male dominated domain, with many investigations finding that men are more likely to be highly involved sport fans than women (for a review, see Wann, Melnick, Russell, & Pease, 2001). However, recent investigations indicate that the number of women sport fans and spectators is on the increase. For instance, the National Football League and National Hockey League estimate that women now constitute roughly 45 percent of their fans (Meyers, 1997; Mihoces, 1998), while women comprise approximately half of major league soccer fans (Hofacre, 1994).

With the recent increase in the number of women entering the male dominated area of sport fandom, researchers have become interested in gender differences in sport fandom as well as differential perceptions of male and female fans. In their work on sport fan motivation, Wann and his colleagues have found several consistent differences in motives expressed by male and female fans (Wann, 1995; Wann, Schrader, & Wilson, 1999). Their research indicates that males are more likely to be motivated by eustress (i.e., positive stress), self-esteem needs, and the aesthetic features of athletic contests. Females are more likely to participate in sport fandom because it provides an opportunity to spend time with their family. Dietz-Uhler, Harrick, End, and Jacquemotte

---

Author info: Correspondence should be sent to: Daniel L. Wann, Department of Psychology, Murray State University, Murray, KY 42071 or to danwann@msumusik.mursuky.edu via Internet.

North American Journal of Psychology, 2001, Vol. 3, No. 2, 183-192.

© NAJP

(2000) also found that female fans were more likely than males to view the activity as an opportunity to be with one's family.

With respect to research on perceptions of male and female fans, Dietz-Uhler, End, Jacquemotte, Bentley, and Hurlbut (2000) asked undergraduates to complete a questionnaire packet assessing their level of identification with their favorite team, level of sexism, and their perceptions of male and female fans. The results revealed that male and female fans were perceived quite differently. For instance, when asked to consider the behaviors of a hypothetical male and female fan, the participants estimated that the male fan engaged in significantly more fan behaviors (e.g., watching sport on television, discussing sport, etc.) than did the female fan. Participants with a high level of team identification and those with a high level of sexism were particularly likely to perceive differences in male and female fans.

While the Dietz-Uhler, End, et al. (2000) work was an important step in understanding perceptions of female sport fans, these authors were simply interested in differential perceptions of fan behaviors as a function of gender. As a result, they did not examine the valence (i.e., positivity or negativity) of males' impressions of female fans. If, as many have suggested, sport is often viewed as a male dominated activity (Blinde, Greendorfer, & Sankner, 1991; Dietz-Uhler, End, et al., 2000; Wann, Schrader, Allison, & McGeorge, 1998), what are the perceptions males possess of female fans, persons who were formerly outsiders for this activity? This question was the focus of the current experiment. Specifically, the current study attempted to determine if males form predominately positive or negative impressions of females who express a high level of interest in sport. Interestingly, using both research and theory, it appears that one could build a fairly sound argument for either impression. On the one hand, in light of the fact that sport has historically been a male dominated institution, it seems reasonable that male fans may reject female fans, viewing them as interfering with their "male bonding" time and violating a well-established societal norm (see Wenner, 1996). This was suggested by Dietz-Uhler, End, et al. (2000, p. 95) who note that males may be threatened by female fans and that they may respond to this threat by distancing themselves from female fans (see also Wann, in press). On the other hand, it is also plausible that male fans hold female fans in high esteem because they share an interest in their choice of pastimes. Research in social psychology reveals that attitude similarity is a powerful determinant of interpersonal attraction (Byrne, Clore, & Smeaton, 1986; Caspi, Herbener, & Ozer, 1992). Further, research investigating the relationship satisfaction of male and female fans supports the important impact of attitude similarity. For instance, Roloff and Solomon (1989) found strong positive correlations between similarity of sport viewing preferences and relationship satisfaction among persons involved in romantic relationships.

The purpose of the current experiment was to test the two competing predictions. Because the research on attitude similarity was so robust (and included empirical support among fans) while the notion that male fans reject female fans was merely speculation, we hypothesized that males with a high interest in sport fandom would give more positive evaluations of women described as having a similar interest than women described as being less than interested in sport. Thus, a two-way interaction involving the participants' level of fandom (high or low) and the female target's level of fandom (high or low) was predicted.

A final variable, the female target's expressed interest in a local sport team, was also investigated. It also seemed that, in addition to her interest in sport in general, a female's interest in a local sport team (or lack there of) may also impact male fans' impressions of her. For instance, an individual who was a fan in general but rejected the home team may be viewed negatively. Unfortunately, as there had yet to be any research in this area with respect to sport fandom, the formulation of direct hypotheses was not possible. Thus, the impact of the female target's interest in a local team was examined as a research question stating "Does a female's level of interest in a local sport team impact male's impressions of her?"

METHOD

Participants and design

Participants were 120 male college students participating in exchange for extra credit in their psychology course (mean age = 22.01 years, *SD* = 4.56). The design was a 2 (Participants' Level of Fandom: high or low) x 2 (Target's Interest in Sport: high or low) x 2 (Target's Interest in the Local Team: high or low) between-subjects factorial. Participants were randomly assigned to the Target's Interest in Sport and Target's Interest in the Local Team conditions.

Procedure and materials

Upon entering the testing session and providing their consent, the respondents were handed a questionnaire packet containing five sections. The respondents were asked to complete the first two sections and then to wait for further instructions. The first section contained demographic items requesting age and gender. The second section contained the Sport Fandom Questionnaire (SFQ). The SFQ contains five Likert-scale format items and is a reliable and valid assessment of one's identification with the role of sport fan (Wann, in press). Response options to the questions ranged from 1 (*low fandom*) to 8 (*high fandom*).

After they had completed the first two sections, the participants were told to read the third section but not to turn to the next page (i.e., the fourth and fifth sections). The third section contained a description of a "fellow college student." The participants were instructed to read the description "slowly and carefully, paying attention to each trait" and to "Use the description to form an impression of this person." The description of the target person contained a list of 18 traits, such as "She once lived in Canada" and "She is a marketing major." Each trait began with "She" or "Her" to explicitly indicate that the target was female. There were four versions of the target description. The first 16 trait descriptors were generic (e.g., "a junior", "is blonde", etc.) and identical for each of the four versions. The 17th trait manipulated the target's interest in sport. Subjects read either "She has a great deal of interest in sports, knows a great deal about the rules, teams, players, and stats, and watches a great deal of sports on TV and in person" or "She has very little interest in sports, knows very few rules, teams, players, or stats, and rarely watches sports on TV or in person." The 18th trait manipulated the target's interest in their university's men's basketball team. Subjects read either "She has a great deal of interest in Murray State Men's Basketball, loves to support the team, wears the team's colors, and loves to watch Murray State Men's Basketball on TV and in person" or "She has very little interest in Murray State Men's Basketball and rarely watches Murray State Men's Basketball on TV and in person." At the bottom of the page, the subjects read "Do not turn to the next page until instructed to do so!"

Once all subjects had carefully read the target's description, they were asked to turn to the last page of the questionnaire packet which contained the final two sections. The participants were explicitly told not to turn back to previous pages while completing this page (i.e., not to refer back to the target's description). The fourth section contained four recall items on the target's traits, two of which were filler items and two of which were manipulation checks. The filler items asked the participants to indicate the target's religious affiliation and the type of vehicle she drove. The manipulation check items asked the participants to indicate whether the target had "very little" or "a great deal of" interest in sports and (separately) Murray State Men's Basketball (subjects were asked to circle the correct response to the manipulation check items). The fifth section of the packet contained 6 items assessing impressions of the target. Participants indicated how friendly, nice, interesting, fun, attractive, and intelligent the target sounded from the description. Responses to the Likert-scale impression items ranged from 1 (*negative impression*, e.g., *not at all friendly*) to 8 (*positive impression*, e.g., *very friendly*). Once the subjects had completed the impression items, the researcher collected the questionnaire packets. Participants were then debriefed and excused from the testing session. The entire procedure lasted about 20 minutes.

RESULTS

All participants correctly stated the target's level of interest in sport and the local team. Consequently, all subjects were included in the following analyses. The five items comprising the SFQ were combined to form a single index of sport fandom (Cronbach's alpha = .92). A median split was performed on the participants' SFQ scores, resulting in two groups: participants with a low level of fandom (*n* = 57, SFQ range = 5 to 27) and participants with a high level of fandom (*n* = 63, SFQ range = 28 to 40). The six impression items were combined to form a single index of overall impression of the target (alpha = .84). Scores on this index were then divided by six to acquire scores that were consistent with the original parameters of each impression item (i.e., scores on the impression index could range from 1.00 to 8.00).

TABLE 1 Means and SDs for Evaluations of the Female Target as a Function of the Participant's Level of Fandom, Target's Level of Fandom, and Target's Interest in the Local Team.

	Participants' level of fandom			
	Low		High	
Target has little interest in sport				
Target has little interest in local team	6.02ab	(.99)	5.41a	(.82)
Target has much interest in local team	6.03ab	(.55)	6.29ab	(.84)
Target has much interest in sport				
Target has little interest in local team	5.33a	(1.66)	6.70b	(.55)
Target has much interest in local team	6.57b	(.89)	6.67b	(.77)

Note: Means with common letters are not significantly different (Newman-Keuls' tests).

The impression index was examined through a 2 (Participants' Level of Fandom: high or low) x 2 (Target's Interest in Sport: high or low) x 2 (Target's Interest in the Local Team: high or low) between-subjects ANOVA (see Table 1 for means and standard deviations). The target's interest in sport main effect was significant, $F(1, 112) = 5.22$, $p < .05$. The target with a high interest in sport ($M = 6.34$, $SD = 1.15$) was rated more positively than the target with little interest in sport ($M = 5.93$, $SD = 0.85$). The target's interest in the local team main effect was also significant, $F(1, 112) = 8.58$, $p < .01$. The target with a high interest in the local team ($M = 6.39$, $SD = 0.79$) was rated more positively than the target with little interest in the local team ($M = 5.89$, $SD = 1.18$). The participants' level of fandom main effect was not significant.

The hypothesized Participants' Level of Fandom by Target's Interest in Sport two-way interaction was significant, $F(1, 112) = 7.31$, $p < .01$. As predicted, the most positive impressions of the target person were reported by participants with a high interest in sport who read that the target had a similarly high interest in sport (*M* = 6.68, *SD* = 0.66). Respondents with a high degree of interest in sport gave less positive ratings to the target with little interest in sport (*M* = 5.83, *SD* = 0.93). Respondents with a low level of fandom also gave less positive evaluations of the target, both when she was described as possessing a high interest in sport (*M* = 5.90, *SD* = 1.48) and when she was presented as having little interest in sport (*M* = 6.03, *SD* = 0.78). Post hoc tests (Newman-Keuls) indicated that the evaluations of the target with a high interest in sport reported by participants with a similarly high interest were significantly more positive than each of the other three conditions, which did not differ from one another. The other two-way interactions were not significant.

The three-way interaction was also significant, $F(1, 112) = 9.82$, $p < .01$. As seen in Table 1, participants with a low level of fandom reported a particularly positive impression of the target with both a high interest in sport and a high interest in the local team. These persons gave particularly negative evaluations to the target with a high interest in sport but little interest in the local team. Moderate evaluations were given to targets with little interest in sport. With respect to participants with a high level of fandom, positive evaluations were given to the target with a high level of interest in sport, regardless of whether she professed an interest in the local team. Fairly positive evaluations were also given to the target reporting little interest in sport but a high interest in the local team. Participants with a high level of fandom gave more negative evaluations to the target described as having little interest in both sport in general and the local team.

DISCUSSION

The hypothesis that males with a high interest in sport would report particularly positive evaluations of a female target with a similarly high interest was supported. Less positive evaluations were expressed by males with a high interest in sport who evaluated a female with little interest in sport as well as participants with little interest in sport (regardless of the target's interest). Thus, contrary to the suggestions offered elsewhere (Dietz-Uhler, End, et al., 2000), it appears that rather than rejecting them, male fans view female fans in a positive light. We have chosen to refer to this phenomenon as the "Something About Mary" effect, based on the popular movie released in 1998. In this film, the producer and director developed the main character, Mary, to be the "perfect" female. While there were many characteristics to her allure to men (e.g., physical attractiveness, wealth, fun-loving nature), a main component of her personality was her high level of interest in sport. She was shown actively participating in sport numerous times and explicitly expressed her interest in watching and following sport even more frequently. Indeed, at one point in the film she asked a male companion, "Hey, do you want to go upstairs and watch Sportscenter?" The film-makers certainly knew what they were doing by portraying the "perfect" woman as possessing an interest in sport, at least with respect to male fans (and recall that most men are fans). The data presented above clearly support the "Something About Mary" effect as male fans in our sample formed particularly positive impressions of our "Mary", i.e., the female target described as possessing a strong interest in sport.

It is important to note that although the current data contradict the suggestions offered by Dietz-Uhler, End and their colleagues (2000), the data described above is still compatible with those presented by these past researchers. For instance, recall that Dietz-Uhler, End, et al., did not assess the valence of their subjects' impressions and perceptions of female fans. Rather, they examined whether or not the impressions of male and female fans differed in some way (e.g., were male fans thought to engage in more intense spectatorship behaviors such as sport viewing). The Dietz-Uhler, End, et al. data indicate that male and female fans are viewed differently. This finding can be integrated with the data described above. Specifically, it may be that male fans do view male and female fans differently with respect to their fanship and spectatorship behaviors. However, different does not necessarily mean negative. Rather, what appears to be occurring is that although female fans are perceived differently, they still are viewed in a positive fashion.

With respect to the research question on the target's interest in a local sport team, the results revealed that her interest did have an impact but, perhaps surprisingly, more so on persons with little interest in sport than on those with a great deal of interest. With respect to those with a high level of interest in sport, higher evaluations were reported for the target with a similarly high interest in sport, regardless of her interest in the local team (the aforementioned "Something About Mary" effect). Although subjects with a high interest in sport also formed somewhat high impressions of the target with little interest in sport but strong interest in the local team, post hoc tests revealed that this effect was not significant. Thus, simply expressing an interest in a local team is not sufficient for fostering the "Something About Mary" effect and the corresponding positive evaluations. Rather, the female must profess an overall interest in sport, not simply an interest in one team (i.e., a fan in general rather than a highly identified fan of just one team).

With respect to the female's interest in a local team on impressions formed by those with little interest in sport, two findings warrant discussion. First, participants with little interest in sport formed significantly less positive impressions of the female target when she was described as having a strong interest in sport but little interest in the local team. The rather negative evaluations of this person may have been due to this target being perceived as a "traitor". That is, although the participants in question had little interest in sport, many of them likely had strong emotional ties to their university. When the target was described as possessing an interest in sport yet rejected their school (i.e., did not support the university's men's basketball team), participants with little interest in sport may have questioned the target's overall loyalty to their alma mater. This line of reasoning does have support both from research on the black sheep effect in social psychology (Marques, Yzerbyt, Leyens, 1988) and work on sport fans in general, who clearly view "traitors" to a local team in a negative fashion (Branscombe, Wann, Noel, & Coleman, 1993).

A second interesting finding on the target's interest in a local team on participants with little interest in sport concerns the particularly positive evaluations these persons reported when the target was presented as possessing both a strong interest in sport and a strong interest in a local team. Based on the previously described importance of similarity in fostering interpersonal attraction, it may have been more logical for these persons to have rejected the target with strong interests in sport and the local team because this target was dissimilar to them on this trait. One may be tempted to use complementarity (Winch, 1958) to explain this phenomenon, suggesting that the males with little interest in fandom were simply attracted to those with the complementary trait of possessing interest in sport. However, such an explanation is inconsistent with the previously discussed finding that those with a high interest in sport but little interest in a local team received the least positive evaluations. An explanation based on complementarity would have required that these evaluations also be positive because this target should also have been perceived as a compliment to the participants with low levels of fandom.

Thus, we are left wondering about the underlying cause of this finding. Two possibilities present themselves. First, perhaps males with little interest in sport are attracted to women with a high interest in sport and a local team because these persons have the potential for enhancing the male's social integration. For instance, these men may perceive some utility in dating women with an interest in sport and a local team because the women may be able to introduce them to social groups that they might not otherwise encounter. The fact that males with little interest in sport fandom reported positive evaluations of the female only when she was described as having a high general interest in sport *and* a local team is consistent with the social integration hypothesis. That is, it would certainly be easier to increase social connections through one's association with a target person if the target were interested in a locally important phenomenon, in this instance, a local sport team. A female fan who did not support the local team would provide much less of an opportunity for integration (e.g., she would not be likely to attend home games, team rallies, etc.).

Secondly, in light of the fact that most males are sport fans (Wann, Melnick, et al., 2000), male nonfans may view female fans as possessing the ability to enhance their acceptance among male peers who are fans. The logic here is that males with little interest in sport may feel uncomfortable in some social situations, such as social gatherings populated by male fans. By connecting and associating with a female fan, they may be able to strategically manage their impression and, in situations where it is warranted, present themselves as a fan through their relationship with a female fan. Self-presentation strategies such as this are quite common among fans (Cialdini et al., 1976; Cialdini, & Richardson, 1980; Leary, 1992; Wann, Royalty, & Roberts, 2000).

In conclusion, because the current investigation was the initial examination of the positivity and negativity of males' perceptions of female sport fans, much remains unknown. Consequently, based on the current data, a number of logical and important suggestions for future research can be offered. For instance, the current study focused on the female target's involvement with a men's sport team. It is unclear whether the current data would generalize to impressions of women who identify with women's team. Also, the current investigation focused solely on males' impression of female fans, leaving uncertain how women view these targets. And finally, the current study targeted involvement with a collegiate team. Certainly, it would be important to replicate the current findings with fans' and targets' involvement with national and professional teams.

REFERENCES

Blinde, E. M., Greendorfer, S. L., & Sankner, R. J. (1991). Differential media coverage of men's and women's intercollegiate basketball: Reflection of gender ideology. *Journal of Sport & Social Issues, 15,* 98-114.

Branscombe, N. R., Wann, D. L., Noel, J. G., & Coleman, J. (1993). In-group or out-group extremity: Importance of the threatened social identity. *Personality and Social Psychology Bulletin, 19,* 381-388.

Byrne, D., Clore, G. L., & Smeaton, G. (1986). The attraction hypothesis: Do similar attitudes attract anything? *Journal of Personality and Social Psychology, 51,* 1167-1170.

Caspi, A., Herbener, E. S., & Ozer, D. J. (1992). Shared experiences and the similarity of personalities: A longitudinal study of married couples. *Journal of Personality and Social Psychology, 62, 281*-291.

Cialdini, R. B., Borden, R. J., Thorne, A., Walker, M. R., Freeman, S., & Sloan, L. R. (1976). Basking in reflected glory: Three (football) field studies. *Journal of Personality and Social Psychology, 34,* 366-375.

Cialdini, R. B., & Richardson, K. D. (1980). Two indirect tactics of image management: Basking and blasting. *Journal of Personality and Social Psychology, 39,* 406-415.

Dietz-Uhler, B., End, C., Jacquemotte, L., Bentley, M., & Hurlbut, V. (2000). Perceptions of male and female sport fans. *International Sports Journal, 4,* 88-97.

Dietz-Uhler, B., Harrick, E. A., End, C., & Jacquemotte, L. (2000). Sex differences in sport fan behavior and reasons for being a sport fan. *Journal of Sport Behavior, 23,* 219-231.

Hofacre, S. (1994). The women's audience in professional indoor soccer. *Sport Marketing Quarterly, 3,* 25-27.

Leary, M. R. (1992). Self-presentational processes in exercise and sport. *Journal of Sport & Exercise Psychology, 14,* 339-351.

Marques, J. M., Yzebyt, V. Y., & Leyens, J. P. (1988). The "black sheep effect": Extremity of judgements towards ingroup members as a function of identification. *European Journal of Social Psychology, 18,* 1-16.

Meyers, B. (1997, August 28). Feminine touches planned but blood and guts remain. *USA Today,* pp. A1, A2.

Mihoces, G. (1998, May 7). Women checking in more as NHL fans. *USA Today,* pp. C1, C2.

Roloff, M. E., & Solomon, D. H. (1989). Sex typing, sports interests, and relational harmony. In L. A. Wenner (Ed.), *Media, sports, and society* (pp. 290-311). Newbury Park, CA: Sage.

Wann, D. L. (1995). Preliminary validation of the Sport Fan Motivation Scale. *Journal of Sport & Social Issues, 19,* 377-396.

Wann, D. L. (in press). Preliminary validation of a measure for assessing identification as a sport fan: The Sport Fandom Questionnaire. *International Journal of Sport Management.*

Wann, D. L., Melnick, M. J., Russell, G. W., & Pease, D. G. (2001). *Sport fans: The psychology and social impact of spectators.* New York: Routledge Press.

Wann, D. L., Royalty, J., & Roberts, A. (2000). The self-presentation of sport fans: Investigating the importance of team identification and self-esteem. *Journal of Sport Behavior, 23,* 198-206.

Wann, D. L., Schrader, M. P., Allison, J. A., & McGeorge, K. K. (1998). The inequitable newspaper coverage of men's and women's athletics at small, medium, and large universities. *Journal of Sport & Social Issues, 22,* 79-87.

Wann, D. L., Schrader, M. P., & Wilson, A. M. (1999). Sport fan motivation: Questionnaire validation, comparisons by sport, and relationship to athletic motivation. *Journal of Sport Behavior, 22,* 114-139.

Wenner, L. A. (1996). The sports bar: Masculinity, alcohol, sports, and the mediation of public space. In S. J. Drucker & G Gumpert (Eds.), *Voices in the street: Explorations in gender, media, and public space* (pp. 71-100). Creskill, NJ: Hampton.

Winch, R. F. (1958). *Mate selection: A study of complementary needs.* New York: Harper & Row.

Parental Disciplinary History, Current Levels of Empathy, and Moral Reasoning in Young Adults

N. L. Lopez
University of Michigan – Ann Arbor

J. L. Bonenberger,
H. G. Schneider
Appalachian State University

The association between parental discipline styles experienced during childhood and levels of empathy and moral reasoning during early adulthood was examined. One hundred and nine undergraduate college students completed two self-report measures of parental discipline styles experienced during childhood. These measures included the *Behavioral Transgression Scale* (BTS) and a modified version of the *Conflict Tactics Scale* (CTSPC). Empathy levels were measured using the *Questionnaire Measure of Emotional Empathy* (QMEE) and moral reasoning was assessed using the *Defining Issues Test* (DIT). Multiple regression analyses were performed using discipline styles as the predictor variables and the overall measures of moral development and empathy as the criterions. Results indicated that gender, use of corporal punishment, and induction were significant predictors of different levels of empathy. Females reported higher levels of emotional empathy than males. Parental use of corporal punishment and low use of induction were associated with low levels of empathy. Minor use of corporal punishment was the only significant predictor of moral reasoning. Young adults who experienced high levels of aggressive punishment during childhood displayed low levels of moral development as measured by the DIT.

Parental disciplinary styles have been moderately associated with the development of empathy and moral thought in children and adolescents (Eisikovits & Sagi, 1982; Grusec & Goodnow, 1994; Hoffman, 1975a, 1975b, 1979; Hoffman & Saltzstein, 1967; Kochanska, 1997; Larzelere, Sather, Schneider, Larson, & Pike, 1996; Larzelere, Schneider, Larson, & Pike, 1998). Most of these studies suggest that children of parents who use inductive discipline styles (such as explaining the effect the child's misbehavior has on others) also display high levels of norm internalization and moral development. However, the relationship between parental use of power assertive techniques and the development of

---

Author info: Correspondence should be sent to: Dr. H. G. Schneider, Psychology Dept., Appalachian State U., Boone, NC 28608

North American Journal of Psychology, 2001, Vol. 3, No. 1, 193-204.
© NAJP

empathy or moral judgment is less consistent. While some researchers have been unable to find an association between power assertion and moral reasoning (Eisikovits & Sagi, 1982; Holstein, 1972), others suggest that power assertion strengthens the efficacy of parental use of induction in increasing norm compliance among toddlers (Larzelere, Sather, Schneider, Larson, & Pike, 1996; Larzelere, Schneider, Larson, & Pike, 1998). This suggests that power assertion may facilitate moral development. The present study attempted to explain these inconsistent findings by refining the assessment and conceptualization of power assertion and differentiating between aggressive and non-aggressive power assertive techniques.

Parental Disciplinary Styles

Most studies divide parental disciplinary interventions into two categories: power assertion and induction (Eisikovits & Sagi, 1982; Hoffman & Saltzstein, 1967; Rich, 1993). Power assertion is a disciplinary method by which parents use their power differential with the child to compel him/her to change his/her behavior (Hoffman, 1975). These interventions include spanking, scolding, withdrawal of privileges, and time outs. In contrast, induction refers to techniques by which the parent ensures that the child understands the consequences of his/her actions on self and others (Hoffman, 1979). These interventions include explaining the consequences of the behavior to the child and explaining the rationale behind the rule. Inductive interventions may contain a power assertive component, but this component is subtle, and designed to obtain the child's attention while not eliciting a strong affective reaction from the child (Hoffman, 1994).

Hoffman (1977) provided a theoretical basis for the role of parental discipline techniques in the development of moral reasoning and empathy. He suggested that the development of moral reasoning is contingent upon the development of norm internalization, which is the ability to suppress inappropriate needs and replace them with socially accepted behaviors without regard for external sanctions. Children develop this ability as they begin to socialize with others, and are also influenced by their disciplinary interactions with their parents. Hoffman (1994) argued that norm internalization is enhanced with the use of discipline styles that elicit an optimal level of arousal to capture the child's attention, but provide the child with minimal external justification for behavior change. When complying with norms, the child is unable to attribute his/her behavior to fear of external sanctions and thus attributes it to an internal attitude change. Children who externalize norm compliance may, as adults, make decisions concerning right and wrong based on fear of punishment or need for reward. According to Kohlberg

(1976), this exemplifies lower stages of moral development. Likewise, people who do not externalize norm compliance during childhood may be more likely to make decisions about right and wrong based on internalized views of justice and equity, and thus have higher levels of moral thought.

Several studies have provided evidence for this theory, suggesting that children of parents who use mostly inductive techniques have higher internalization of norms, more guilt over antisocial behavior, and higher levels of empathy, altruism, and moral reasoning (Eisikovits & Sagi, 1982; Hoffman, 1969, 1975a, 1975b, 1979; Hoffman & Saltzstein, 1967; Krevans & Gibbs, 1996). However, the relationship between induction, moral thought and norm internalization appears to be more complex. Some studies indicate that the child's temperament and the child's acceptance or rejection of the discipline technique moderate the effect of induction and power assertion on the development of moral thought (Grusec & Goodnow, 1994; Kochanska, 1997). Others also suggest that the use of induction alone is not effective, and that it is associated with a variety of negative consequences (Larzelere et al., 1996, 1998).

Larzelere et al. (1998) studied the effectiveness of parental reasoning in delaying toddler misbehavior. Reasoning was most effective when combined with punishment when the children were 2-3 years old and during a 20-month follow- up assessment. Most importantly, children whose mothers reported using reasoning without punishment had the largest increase in disruptive behavior from 2 to 4 years of age. Larzelere et al. (1996) also found that reasoning combined with mild punishment was the most effective intervention in delaying toddler misbehavior when compared with using punishment or reasoning alone. However, the combination of non-aggressive punishment and induction was more consistently associated with a delay in toddler misbehavior than the combination of aggressive punishment and induction. Although these studies only addressed norm compliance, it can be argued that toddler misbehavior is related to norm internalization, and thus reasoning alone may not be an effective method of facilitating the development of norm internalization.

The findings of Larzelere et al. (1996, 1998) are consistent with Hoffman's (1994) theory. Hoffman argues that parental interventions must produce some physiological arousal in order to gain the child's attention, but he also indicates that parental interventions that elicit high levels of anxiety, fear, and anger interfere with the development of empathy, guilt and moral reasoning. Elevated levels of anxiety and fear re-direct the child's attention towards external sanctions, and may prevent the child from understanding the impact of his/her behavior on others. Likewise, too little arousal may result in the child ignoring the parent and disregarding the parent's inductive message.

Although the level of aggression within power assertive interventions may play an important role in determining the influence of induction on the development of moral reasoning and empathy, most studies have not considered this factor (Eisikovits & Sagi, 1982; Holstein, 1972). The purpose of this study was to analyze the possible role of parental induction and power assertion on the development of empathy and moral reasoning while distinguishing between aggressive and non-aggressive types of power assertion. Children's emotional reaction to highly aggressive discipline may interfere with norm internalization even when these interventions are accompanied by inductive interactions. Thus, we hypothesized that young adults who experienced high levels of aggressive discipline styles during childhood would have lower levels of empathy and moral development when compared to young adults who did not experience aggressive discipline interventions during childhood. It was also expected that inductive discipline, while controlling for power assertion, would be associated with high levels of moral development and empathy.

METHOD

Design and Participants

One hundred and two undergraduate students from a liberal arts university in the southeastern part of the United States volunteered for the study. Each of the volunteers completed a testing package consisting of four measures presented to them in a random order. Participants completed a consent form and were tested in small groups. The sample consisted of 63 females and 39 males (mean age = 19.67 years, *SD* = 2.01). Ninety five percent of the subjects were Caucasian (*n*=97). Two multiple regression analyses were used with disciplinary styles and demographic information as predictor variables, and empathy and moral development as the criterion variables.

Measures

Moral Development. Levels of moral judgment were measured using the *Defining Issues Test* (DIT; Rest, 1986). The DIT assesses the cognitive processes used during the analysis of social-moral problems. Participants are presented with a series of moral dilemmas and are asked to reach a decision concerning the most appropriate solution to each situation. Participants are then asked to rate a series of factors that could have influenced their decision. The "P" Index of this scale was used to measure Principled Morality. High scores on this index indicate a tendency to make moral decisions based on consideration of rational thought concerning laws and norms, such as whether these norms reflect the general will of society and guarantee basic rights (freedom, life, happiness) to all members of society, and whether such rules are congruent with the principles of justice and equity (Rest, 1979). The DIT has high internal consistency, with Cronbach's alphas for most sub-scales in the high .70s (Davidson and Robbins, 1978), and moderate test-retest reliability (r = .77).

Parental Discipline Styles. In order to adequately assess a broad range of disciplinary interventions, scales were selected from two different instruments measuring parental disciplinary choices. Three of the sub-scales from *Conflict Tactics Scale* (CTSPC; Straus, 1997), one of the most reliable tools for the assessment of domestic violence, were used to assess levels of aggressive parental discipline. This scale is a 35-item questionnaire designed to assess the disciplinary methods of families and the level of aggression present within families. The following scales from the CTSPC were included: Psychological Aggression (threats of violence, shouts, screams, insults, etc), Minor Assault (corporal punishment: spanking, hitting, slapping, etc.), and Severe Assault (hitting the child with objects, hitting or punching the child on his/her face, kicking the child, etc). The CTSPC has adequate test-retest reliability ($r = .80$), but has lower internal consistency because of the heterogeneity of the items and scales (Amato, 1991). The validity of the CTSPC has been assessed by correlating the scale with factors associated with corporal punishment such as age of parent, age of child, race, and gender of parent.

To assess non-physical interventions, the Induction and Non-aggressive Power Assertion sub-scales from the Behavior Transgression Scale were used (BTS; Lopez, Dula, & Schneider, 1999). The BTS consists of 20 hypothetical situations representing 4 types of pre-adolescent and adolescent age behavioral transgressions (aggressive and non-aggressive safety violations, and aggressive and non-aggressive social violations). The items contain excellent inter-rater reliability with $r > .90$ (Lopez, Dula & Schneider, 1999). Examples of these transgressions include "pointing a gun at someone" (aggressive safety violation), "Using profanity towards a teacher" (aggressive social violation), "trying to get something out of the oven without gloves" (non-aggressive safety violation), and "Leaving the refrigerator's door open" (non-aggressive social violation). Subjects were presented with these items and asked to state how likely their parents were to use different interventions when addressing each transgression when they were between 8 and 13 years old. The BTS provides scores for overall use of Induction (explaining the consequences of his/her actions on self or others), Aggressive Power Assertion (spanking, verbal aggression), and Non-Aggressive Power

Assertion (withdrawal of privileges, time out). The aggressive power assertion subscale was not used because it presented a significant correlation with the CTSPC minor assault subscale (r = .72). The CTSP was selected over the BTS *Aggressive Power Assertion Scale* because of the CTSP's established psychometric proprieties, and to facilitate comparisons with previous studies of parental discipline. Furthermore, including both subscales in the regression models could result in an overestimation of these variables' regression coefficients and an increase of the standard error of the estimate.

Empathy. The *Questionnaire Measure of Emotional Empathy* (QMEE; Mehrabian & Epstein, 1972) was used to measure levels of empathy. This 33-item questionnaire has good internal consistency (split-half reliability of .84). Participants are asked to rate on a 9-point Likert scale their level of agreement with a number of statements, such as "It makes me sad to see a lonely stranger in a group" and "I get really angry when I see someone being ill-treated." The QMEE is negatively correlated with measures of aggressive behavior, and positively correlated with measures of helping behavior. The QMEE also shows good discriminant validity, and it is not contaminated by social desirability.

RESULTS

The first multiple regression was performed to determine which variables predict empathy as measured by the QMEE. The predictor variables consisted of a series of demographic factors (gender, age, and person in charge of discipline), and parental use of induction (BTS), non-aggressive power assertion (BTS), major and minor physical discipline (CTSPC), and psychological aggression (CTSPC). Gender was coded into the regression as a dichotomous nominal variable (female = 0; male = 1). The person in charge of discipline was coded as a polychotomous nominal variable. Participants were asked to indicate who was in charge of enforcing discipline during their childhood. Possible answers included mother, father, both, or other. The option "other" was not selected by anyone and was not used. This regression model used "both" as the omitted, reference variable.

The regression model was significant F (9,92) = 4.27, $p<.01$. The multiple correlation coefficient was .54, which accounted for 29% of the variance (adjusted R^2 was .23). Table 1 contains the summary of this regression model. Three variables were significantly related to empathy. Gender was significantly related to empathy (b = -18.72, p = .01), with females averaging 18.72 points higher than males on the QMEE. Parental use of induction was also a significant predictor of empathy levels ($b = .27, p < .05$). Finally, use of minor corporal punishment was

TABLE 1 Summary of Regression Analysis for Variables Predicting Empathy

Predictor Variable	*B*	β	Bivariate Correlation	Partial Correlation
Gender	-18.72	-.30	-.38	-.32**
Age	2.38	.16	.11	.18
Disciplined by Mother	-6.12	-.09	-.09	-.09
Disciplined by Father	-2.74	-.03	-.07	-.03
Induction	.27	.24	.28	.24*
Non-agg. Power Assert.	-.001	-.01	-.01	-.01
Psych. Aggression	.20	.15	-.11	.11
Phys. Aggress.- Minor	-.39	-.23	-.30	-.23*
Phys. Aggress.- Severe	-1.0	-.22	-.16	-.15

* $p < .05$ ** $p < .01$

negatively associated with empathy scores ($b = -.39$, $p < .05$). The standardized coefficients of this model indicate that gender was the strongest predictor of empathy, followed by induction and the minor use of corporal punishment. No other discipline style was significantly related to empathy.

A second multiple regression analysis was performed to predict levels of moral development based on parental disciplinary styles and demographic variables. Empathy was also included in this model as a

TABLE 2 Summary of Regression Analysis for Variables Predicting Principled Morality

Predictor Variable	*B*	β	Bivariate Correlation	Partial Correlation
Gender	-.36	-.01	-.12	.01
Age	-.08	.01	.09	.01
Disciplined by Mother	2.48	.08	.08	-.08
Disciplined by Father	2.29	.04	.02	.05
Induction	-.04	-.09	-.03	-.09
Non-agg. Power Assert.	-.06	-.13	-.20	-.14
Psych. Aggression	-.02	.03	-.06	-.05
Phys. Aggress.-Minor	-.27	-.33	-.37	-.30**
Phys. Aggress.-Severe	.17	.07	-.06	-.05
Empathy	.10	.22	.28	.20*

* $p < .05$ ** $p < .01$

predictor of moral judgment. The regression model was significant, F (9,92) = 2.54, $p<.01$. The multiple correlation coefficient was .47, which indicated that approximately 22% of the variance was accounted for by the model (adjusted R^2 was .13). Table 2 contains the summary of the regression model. Two variables were significant predictors of moral judgment. Minor use of physical discipline was significantly related to Principled Morality ($b = -.27$, $p <.01$). Empathy was also a significant predictor of moral judgment ($b = .10$, $p < .05$). While empathy and induction were significantly related, induction was not significantly correlated with moral judgment.

DISCUSSION

The purpose of this study was to determine if disciplinary styles predict levels of moral reasoning and emotional empathy. Both regression models were significant, indicating that parental discipline styles were related to levels of empathy and moral reasoning. Minor use of corporal punishment, independent of use of induction and a variety of other variables, were significant predictors of low levels of emotional empathy. These results are consistent with Hoffman's (1994) theory, which predicts that the use of harsh discipline prevents the child from focusing his attention on the consequences of his/her actions on others. By only focusing on personal consequences, the child may be unable to experience feelings of empathy and guilt, and thus may create a system of values that limit him/her from experiencing and expressing empathy.

Parental use of induction was also positively related to high levels of empathy. This is consistent with Krevans & Gibbs' (1996) findings, which indicate that parents who use high levels of induction also have children who experience higher levels of empathy and pro-social behavior.

Empathy was also significantly related to gender, with females endorsing higher levels of empathy than males; other studies have found similar results (Sinclair & Bourne, 1998; Trobst, Collins, & Embree, 1994). However, the nature of the relationship between gender and empathy appears to be related to the type of measurement used during the study. Eisenberg and Lennon (1983; 1987) performed extensive meta-analyses of the literature of gender and empathy and concluded that while females report higher levels of empathy than males during self-report tests, these gender differences are not significant when empathy is assessed using physiological measures and/or unobtrusive observations.

While non-aggressive interventions were unrelated to moral reasoning, aggressive power assertion was a moderate but significant predictor of low levels of principled morality. Parents who use aggressive power assertive techniques may preclude the development of an internalized set of values, which can result in an over-reliance on external sanctions when making decisions about moral issues. The use of aggressive discipline styles may also elicit high levels of emotional arousal that may prevent the processing of the inductive message present within the discipline encounter. However, the small size of the correlation between aggressive power assertion and moral reasoning found in this study suggests that many other factors may moderate the relationship between these variables.

An alternative explanation of these results postulates an interaction between a child's characteristics and parental disciplinary choice. Children with lower internalization of norms may misbehave more, which may elicit higher levels of aggressive discipline. Although this explanation seems logical, Lopez, Dula, & Schneider (1999) studied the effects of different levels of behavioral transgressions on parental disciplinary responses, and found that parents increase the use of all forms of discipline interventions (not only corporal punishment) when confronted with more severe violations. If children with low moral development simply require higher levels of discipline, a relationship between high levels of all types of discipline interventions and moral reasoning would be expected. This was not supported, since only minor use of corporal punishment was significantly related to moral reasoning. Although the findings of Lopez et al. (1999) appear to rule out this alternative interpretation, more controlled studies are needed to test this hypothesis.

Surprisingly, severe corporal punishment was not a predictor of either empathy or moral development, primarily because of the very low levels of severe punishment endorsed by the participants, which significantly reduced the variance of this variable (67% of the sample received a zero on the severe punishment subscale). Severe physical punishment may have been under-reported, but it is more likely that the sample actually experienced low levels of this type of punishment during childhood. The current sample contained parents that were highly educated and less likely to use severe corporal punishment.

The appropriateness and usefulness of non-abusive physical discipline are still being debated. This study suggests that aggressive parental interventions are related to some negative outcomes. The foundation of our interpretation is that the use of aggressive punishment may create a state of hyper-arousal that prevents the child from attending to the inductive message from the parent. However, several studies indicate that the effect of physical discipline varies by culture (Deater-Deckard & Dodge, 1997). It is possible, that the experience of high levels of arousal may not be related to the type of discipline used but to the child's temperament, perception of the disciplinary intervention, or his/her culture. Aggressive parental interventions may not elicit strong emotional reactions among children living in specific environmental contexts or cultures. It is also possible that children with specific temperaments may require different parenting styles in order to facilitate moral development. Kochanska (1997) found that gentle discipline was associated with moral development among toddlers with fearful temperament but not among toddlers with fearless temperament.

Future studies should continue to assess the role of temperament and other parental variables in mediating the relationship between aggressive power assertion and the development of empathy and moral reasoning. Also, it is important to perform longitudinal studies that analyze the effects that children with low levels of empathy and moral development have on their parent's choice of disciplinary techniques. Finally, longitudinal studies should directly assess the efficacy of combining induction with different types of power assertive techniques in the development of norm internalization, moral reasoning, and empathy.

REFERENCES

Amato, P. R. (1991). Psychological distress and the recall of childhood family characteristics. *Journal of Marriage and the Family, 53,* 1011-1019.

Bryan, J. W., & Freed F. W. (1982). Corporal punishment: Normative data and sociological and psychological correlates in a community college population. *Journal of Youth and Adolescence, 11,* 77-87.

Cantron, T. F., & Masters, J. C. (1993). Mother's and children's conceptualization of corporal punishment. *Child Development, 64,* 1815-1828.

Davison, M., & Robbins, S. (1978). The reliability and validity of objective indices of moral development. *Applied Psychological Measurements, 2,* 397-403.

Deater-Deckard, K., & Dodge, K. A. (1997). Externalizing behavior problems and discipline revisited: Nonlinear effects and variation by culture, context, and gender. *Psychological Inquiry, 8,* 161-175

Eisikovits, Z., & Sagi, A. (1982). Moral development and discipline encounter in delinquent and non-delinquent adolescents. *Journal of Youth and Adolescence, 11,* 217-130

Eisenberg, N., & Lennon, R. (1983) Sex differences in empathy and related capacities. *Psychological Bulletin , 94,* 100-131.

Grusec, J. E., & Goodnow, J. J. (1994). Impact of parental discipline methods on the child's internalization of values: A reconceptualization of current view points. *Developmental Psychology, 30,* 4-19.

Hoffman, M. L. (1975a). Altruistic behavior and the parent-child relationship. *Journal of Personality and Social Development, 31,* 937-943.

Hoffman, M. L. (1975b.). Moral internalization, parent power, and the nature of parent-child interaction. *Developmental Psychology, 11*, 228-239.

Hoffman, M. L. (1979). Development of though, feeling and behavior. *American Psychology*, October, 958-965.

Hoffman, M. L. (1994). Discipline and internalization. *Developmental Psychology, 30*, 26-28.

Hoffman, M. L., Saltztein, H. D. (1967). Parent discipline and the child moral development. *Journal of Personality and Social Psychology, 5*, 45-57.

Holden, G. W., Coleman, S. M., & Schmidt, K, L. (1995). Why 3-year-old children get spanked: Parent and child determinants as reported by college-educated mothers. *Merill-Palmer Quarterly, 41*, 431-452.

Holstein, C. B. (1972). The relationship of children's moral judgment levels to that of their parents and to communication patterns in the family. In R. C. Smart & M. S. Smart (Eds.), *Readings in child development and relationships*, 484-494. New York: Macmilan.

Kohlberg, L. (1976). Moral stages and moralization: Cognitive-developmental approach. In T. Lickona (Ed.), *Moral Development and behavior: Theory, research, and social issues*. New York: Holt, Rinehart & Winston.

Kochanska, G. (1997). Multiple pathways to conscience for children with different temperaments from toddlerhood to age 5. *Developmental Psychology, 2*, 228-246.

Kravens, J., & Gibbs, J. C. (1996). Parents use of inductive discipline: Relations to children's empathy and prosocial behavior. *Child Development, 67*, 3263-3277.

Larzelere, R. E., Sather, P. R., Schneider, W. N., Larson, D. B., & Pike, P. L. (1998). Punishment enhances reasoning's effectiveness as a disciplinary response to toddlers. *Journal of Marriage and the Family, 60*, 388-403.

Larzelere, R. E., Schneider, W. N., Larson, D. B., & Pike, P. L. (1996). The effects of discipline responses in delaying toddler misbehavior recurrences. *Child and Family Behavior Therapy, 18*, 35-57.

Lennon, R., & Eisenberg, N. (1987) *Gender and age differences in empathy and sympathy*. New York, NY, USA: Cambridge University Press.

Lopez, N. L., Schneider, H. G. & Dula, C. S. (1999). *The behavioral transgression scale: preliminary item development*. Presented at the annual conference of the Southeastern Psychological Association, Savannah, Georgia. Available from Dr. H. G. Schneider, Appalachian State University, Department of Psychology, Boone, NC, 28608.

Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality, 40*, 525-543

Parikh, B. (1980). Development of moral judgment and its relation to family environmental factors in Indian and American families. *Child Development, 51*, 1031-1039.

Rest, J. R. (1979). *Development in judging moral issues*. Minneapolis: University of Minnesota Press

Rest, J. R. (1986). *Moral Development Advances in Theory and Research*. New York: Praeger Press.

Rich, J. M. (1993). Discipline and moral development. *The High School Journal*, Dec/Jan, 139-144.

Salzinger, S., Feldman, R., Hammer, M. & Rozario, M. (1993). The effects of physical abuse on children's social relationships. *Child Development, 64*, 169-187.

Sinclair, H. C., & Bourne, L. E. (1998). Cycle of blame or just world: Effects of legal verdicts on gender patterns in rape-myth acceptance and victim empathy. *Psychology of Women_Quarterly, 22*, 575-588

Straus. M. (1991). Discipline and deviance: Physical punishment of children and other crimes in adulthood. *Social Problems, 38*, 133-152.

Straus, M. (1994). Corporal punishment of adolescents by parents a risk factor in the epidemiology of depression, suicide, alcohol abuse, child abuse, and wife beating. *Adolescence, 29*, 543-561.

Straus, M., Hamby, S., Finkelhor, D., Moore, D., & Desmond, R. (1997). Identification of child maltreatment with the Parent-Child Conflict Tactics Scale: Development and psychometric date for a national sample of American Parents. *Child Abuse and Neglect, 22*, 249-270.

Trobst, K. K., Collins, R. L., & Embree, J. M. (1994). The role of emotion in social support provision: Gender, empathy, and expression of distress. *Journal of Social and Personal_Relations, 11*, 45-62.

Walker, L. J., & Taylor, J. H. (1991). Family interactions and the development of moral reasoning. *Child Development, 62*, 264-283.

An Interview with Russell Eisenman

Russell Eisenman
University of Texas - Pan American

(interviewed on behalf of NAJP by)
Michael F. Shaughnessy
Eastern New Mexico University

NAJP: What are you currently working on/researching/writing ?

RE: I am doing various things in crime and creativity. I am looking at personality characteristics of prisoners, to see if personality traits associated with creativity—such as openness, fluency, etc.—are associated with reduced recidivism. There is some reason to think that more creative thinking in prisoners should lead to reduced recidivism. The typical prisoner is the opposite of the creative person. The prisoners I have worked with tend to be rigid, inflexible, not very fluent, unable to produce many alternatives, etc. They think and behave rigidly with simple, anti-social rules that govern their behavior. For example, many think if you show disrespect to them they should respond with violence, perhaps even murder.

They do not have flexible ways for dealing with what they identify as disrespect. They do not even have flexible ways of seeing that what they call "disrespect" might really be something else, like an accidental behavior on the other person's part. I have done two studies that show that most prisoners are low in creativity, relative to the general population. This goes against the romanticized view of the creative criminal smashing out against the rigid restrictions of society. I think the romanticized view is nonsense. There are a few creative prisoners, and I have identified them in one study I did. They are the minority of the prison population. These creative prisoners, like other prisoners who show some small amount of creativity, mainly use their creativity in the area of crime. You could say that is their specialization, their lifestyle. So, from a societal standpoint, what we want are people who will use their creativity to go away from anti-social criminal activities and become more pro-social. A tough task indeed, to achieve. But, perhaps it can be done and we have to try. The benefits of doing it are great, and I think some prisoners can be reformed. Perhaps most cannot, but many do

---

Author info: Correspondence should be sent to: Dr. Michael Shaughnessy, Psychology Dept., Eastern New Mexico U., Portales, NM 88130.

North American Journal of Psychology, 2001, Vol. 3, No. 2, 205-216.
© NAJP

not want to be criminals, but they have no idea how one does that. So, with training/therapy and help in society, such as job training, drug and alcohol treatment, etc., quite a few could be turned into productive citizens.

Of course, an interesting question is this: Could it be that I find little creativity because I am dealing with the ones who got caught? Perhaps the creative criminals avoid capture—at least most of the time—so they do not get studied. This is a good point. I do not think it is mainly true, although it probably has some truth to it. There may be a few who fit that description of a creative criminal who avoids capture. Many of the prisoners I worked with had not been caught for long periods of time, but they keep committing crimes, eventually get caught, get out and commit more crimes, are not caught for awhile, then finally get caught again, and so on. But, unless we could find creative criminals outside of prison and study them, we cannot really answer the question of whether or not the ones in prison are there because they were not creative enough to evade capture.

A different area I am studying has to do with my interests in multiculturalism. I am looking at African-American and White School children in the first four grades, and how they are referred by their teachers for special education and for physical disability services. By physical disability I do not mean the things that people usually think of, such as being unable to walk, but rather anything that is a physical problem, such as poor eyesight or poor hearing. My findings indicate that the African-American children are much more likely to be seen as needing services for academic and for physical deficiencies. The big question is "What does this mean?" It could mean only that teachers perceive the African American children as needing these services, even if they do not. But, I suspect a lot of it is due to African American children needing services in the two areas: academic and physical. Perhaps many of the prenatal problems of African American women cause their children to have various handicaps. If we can identify these handicaps at an early age we can make tremendous improvements, and help African American children have a fair chance in our society.

NAJP: What do you see as the main issues currently facing clinical psychology?

RE: There are several changes in clinical psychology from the time I was trained. One of the biggest is the recent realization of the importance of genetics. Professions seem to go in cycles. When I was trained I learned that there used to be naïve genetic assumptions, but that behavior seems largely learned. Now, learning is taking something of a back seat to all the genetic findings. Much of what we do seems to have a strong genetic component. Of course, genetics and the environment go hand in hand. We often make artificial attempts to separate them. But genetics is much stronger than many realized. Along the same continuum, there is the field of evolutionary psychology, also known as sociobiology. When I first read a little about how social events could be explained biologically and evolutionarily, I thought it was mostly nonsense. The field seemed to rely on wild speculation. But, as I hear about and read more in the field, a lot of it makes sense. Many of the findings and predictions made by evolutionary psychology make sense to me. Men value youth, beauty and health in choosing a female partner, but women value more the power and status of the male partner; as men get older they are attracted to younger women; and other findings that are explainable on an evolutionary basis, with the main desire being to transmit one's genes into the future. Clinical psychology needs to take into account the new findings in genetics and in evolutionary psychology and sociobiology.

From an applied standpoint clinical psychology needs to deal with managed care. Managed care has taken the power away from the therapist and put it in the hands of someone who is more concerned with making money than with providing proper treatment. There are many horror stories. Clinicians that I know tell me that managed care has reduced their caseloads. They cannot make money like they used to because managed care attracts people who formerly would have come to the clinicians. This is an important issue, having practical and ethical problems.

NAJP: How can we best train clinical psychologists?

RE: We need to train clinical psychologists to know about issues such as genetics, evolutionary psychology, multiculturalism, and other emerging areas, as well as the standard fare. All of education should be more interdisciplinary than it is. I have tried to learn much about areas such as sociology, anthropology, political science, philosophy and other areas that relate to understanding people. Clinical psychology focuses too much on pathology and not enough on normality and creativity. If you think of a continuum with pathology at one end and creativity at the other end, clinical psychology has always focused too much on the pathology end. Creativity is seldom seen as being part of clinical psychology or of social psychology. Yet, it fits importantly into both areas of understanding. A sad fact is that there are more and more Ph.D.s turned out each year, yet there are fewer and fewer academic jobs. Training should prepare students for other areas in addition to being an academic. To take only one possible example, as we have more and more people incarcerated, there is increasing need for psychologists to work in the forensic area. Not just in prisons, but working with courts, judges, district attorneys and the like. By the way, it is not totally bad that we put more and more people in prison. We lead the world, but perhaps we should, given all the crime we have. Perhaps we need a more creative look at crime and ask why we have so much in this nation, and what can be done about it. In the mean time, lots of bad people need to be locked up. One exception to this is the War on Drugs, which has partially been responsible for incarcerating so many. All too often it is the not the big time drug dealer who goes to prison - often for a long time, often where the law mandates a long sentence and gives the judge no discretion—but lower level dealers and user. This is crazy and inhumane. The main reason that the incarceration rate for women has increased so much is the imprisonment of so many women for drug crimes. I used to think that as women worked more that they had more opportunities to commit crime, and that is why so many more were going to prison. But, that is not the main reason. The main reason has been the strict enforcement of drug laws. Both men and women are being put into prison at remarkable rates due to our War on Drugs, which does not seem to work. It is like the "Just Say No" program. Research so far shows that, for the most part, it does not work. But, it sounds good. Cops like it. People who don't know any better like it. It looks good on the six o'clock news show to have friendly cops with nice looking elementary school kids, and the kids screaming "Just say no to drugs." But if it does not work, it is a waste of resources.

NAJP: You have written about the relationship between crime and creativity. Why does this relationship exist and is it just a spurious one?

RE: This question is probably dealing with more than one issue. For one thing, there is a negative relationship between prisoners and creativity in my research. That is, prisoners tend to be low in creativity. Why? Because they come from horrible family backgrounds where they suffered physical, psychological or sexual abuse. They grow up hating authority. They tend to have low intelligence, as measured by various IQ tests. They have little toleration of sitting still in school and learning. They end up as having no real job skills and no desire to work for anyone else and conform to job demands. Perhaps many of them have Attention Deficit Disorder with Hyperactivity. They certainly act as if they do. Or, perhaps they are anti-social personalities and conduct disorders (psychopaths in the making), and will not tolerate anything that does not provide immediate gratification. Anyway, none of this makes for the kind of person who will get the education to learn about ideas and things, and to later come up with creative, innovative ways. Instead, it leads to an angry, authority-hating person who cannot do much in life except work at low-level, low-paying jobs, or go into crime and make more money. For sex offenders, many of their problems also seem traceable to horrible family rearing experiences, often being sexually abused early in life. They go on to do to others what was done to them. Being abused does not usually result in empathy, but instead, in wanting to hurt others as one was hurt.

NAJP: Do crime and creativity sometimes go together?

RE: Yes! Sometimes there are creative criminals, people who use their creativity to further their criminal goals. Many psychopaths throughout history have been successful in life because they used their creative thinking for anti-social purposes. Others may not be psychopaths, but they do psychopathic–type things and profit via creative thinking.

Take Clifford Irving, who fakes a Howard Hughes biography. His manuscript took in McGraw Hill, who paid him a six-figure advance. Had not Howard Hughes reversed his stance of not commenting publicly on anything, and denounced the book as a fraud, Irving would have gotten away with it. As it was, he served a short prison sentence, and came out and wrote a book about his fraud, which became a best seller. I think many public figures, and also people we never hear about, differ on a continuum from having no psychopathic tendency to being your full-blown psychopath (the anti-social personality disorder). Their psychopathic tendencies, especially if supplemented with creative thinking and behaving, help them succeed in life.

Long ago, I concluded that Bill Clinton was way above average on the psychopathic continuum. So I was not surprised at all by the scandals that emerged. I am not saying he is a psychopath, only that he is more so than average. He got caught, but he also got to be President of the United States - not a bad achievement. Perhaps he used psychopathic and creative tendencies to further his goals.

NAJP: Is there true rehabilitation for certain criminals?

RE: I think so. Working with prisoners in the California Youth Authority, I am impressed with how many seem to be beyond help, even if they are only 16 years old. They have had a lifetime of crime, so far, and they do not seem motivated to change. All they know, or care to know, it seems, is how to be the kind of criminal they are: drug dealer, burglar, sex offender, or whatever. However, I would estimate that about 15 % can be helped with intensive psychotherapy, job training, drug and alcohol treatment. Ideally they should continue to get help after they leave the prison.

This last one is typically neglected, but since so many go back to the same environment, there is a great need for post-confinement services. Some might think, why waste your time if 85% will not change? The answer is that there is a great social advantage to turning around 15 % (or whatever it is; I confess my ignorance of the true number who can be helped). If 15% are now no longer robbing, murdering, or raping people, that is a great benefit to society. The problem is that we cannot tell which ones belong to the small number who can be helped and which belong to the probably larger number who cannot be helped. We can sometimes tell, as when a prisoner continues to be very anti-social, even in prison, but we cannot really tell most of the time. So we have to have a broad-based attempt to help as many as possible, so that we reach the ones who truly can be changed.

NAJP: As we enter the year 2001 and the new millennium, what are the challenges facing psychology?

RE: The challenges facing psychology are many. The ones I would emphasize are: how do we deal with such areas as violence, drugs (including alcohol, which is a powerful drug), gangs, different racial and ethnic groups in the country, the lack of achievement motivation in many youth, the lack of parenting by many parents, the increasing reality of the single parent home, and probably many things that I have left out.

Another question is what kind of research do we do and support. Philip Zimbardo, for example, once described himself as a generalist and championed that viewpoint, yet said that the kinds of people who get faculty jobs at universities are the specialists. Perhaps psychology should reconsider the value of knowing more and more about less and less. Another challenge for psychology, totally unrecognized by most, is age discrimination. I have faced this in recent years as I have tried to obtain more suitable academic positions. It is an issue for the entire country, as older people are often laid off from high level as well as low level jobs, to be replaced by younger, cheaper people. In China, age is thought to lead to wisdom, and elders are given much respect. Here in the U.S., there is prejudice against older people, even by psychologists, who should know better. It is pervasive throughout the entire culture. In our movies and television shows, we are constantly bombarded with youth, and this is what most people have come to prefer. They show that same preference when they are in position of authority and can hire someone. Other prejudices as well —racial, ethnic, gender— make our society a lesser place than what it could be.

NAJP: We are increasingly confronted with HMO's and their insistence on medication as opposed to therapy. What will this perspective do to psychology and managed health care?

RE: This will have a horrible effect on psychology and will lead to managed care giving medication too often and psychotherapy too little. HMO's are great in one sense, in that they provide services inexpensively and thus help many who otherwise might not be helped. But, they also do terrible things sometimes, with an emphasis on the bottom line, dollars. Treatment cannot be based on which method will be the least expensive. We have the best and worst kind of capitalism here.

NAJP: Depression remains problematic in spite of Prozac and other drugs. To what do you attribute this dilemma?

RE: It would seem to me that the drugs do not take care of all of the aspects of depression. Perhaps depression is a multi-dimensional concept, rather than one thing. It could also be that psychotherapy is needed along with the drug treatment. People need others to talk to, to get a better understanding of their current situation. Even if one were bio-chemically inclined toward depression, and that was reduced or cured by a drug, there might still be a need to sort things out via counseling, psychotherapy or other "talking cures."

NAJP: Stress in modern life is epidemic. How can psychology help?

RE: Psychology can be of great help here. Much of stress is self induced, or at least made worse by what people think, as Albert Ellis would say. What people tell themselves often makes things worse for them, or is the basic cause of their problems.

My dentist once complimented me on how well I responded to the various dental treatments, and asked if I was using something from psychology. The basic answer was yes, stress reduction. I try not to focus on what the dentist is doing, I tried to focus on something else, such as staring at an object in the room. I try to think of pleasant things and I work on my fears when they occur by telling myself (thinking) that things are not as bad as I am perceiving them, and they will soon be over. All this is basic stress reduction. Psychology provides many insights into how people think, and how we can make things better or worse for ourselves. I constantly talk to people who are non-psychologists, who create stress for themselves by what they think about, or who fail to get out of stress when they could. This is not to say that we should rationalize away true problems. But we should face them rationally, and that often leads to less stress. Other things, such as taking time out from our busy world to relax, or engage in sports or exercise can reduce stress. Another stress reducing method is to figure out what would reduce the stress.

For example if you are nervous about a presentation, you can rehearse it until you have it down well. Also, you can practice the presentation in front of others, to reduce the stress you will feel when the real day comes, and you have to present. Psychologists know a lot, and how to deal with stress is one of the things we are good at.

NAJP: What are your most important writings ?

RE: I write on lots of different topics so there are many different ones that have been influential, or which I am proud of. I will hit some of the highlights here, neglecting others that might be equally valuable. In two studies (Eisenman, 1992a, 1999) I showed that prisoners are low in creativity. In one study (Eisenman, 1999a) I found that psychotic prisoners were lower in creativity than conduct disordered prisoners. These studies go against the popular ideas that criminals are creative, or that creativity is enhanced by psychosis. Much of my work has shown that commonly held ideas are incorrect. Several years ago, there was the popular notion among researchers that psychopaths do not learn. This would seem to explain why they are they way they are.

But, Bernard and Eisenman (1967) showed that not only do psychopaths learn better than the control group, but also contrary to established beliefs among researchers, the psychopaths learned better for social praise than they did for money. Perhaps this study and others like it helped reduce the misconceptions about psychopaths. It is possible that they have some learning deficits in certain areas, but to say that overall, psychopaths cannot learn is incorrect.

Many researchers think that you cannot change sexual behavior. You can change sexual attitudes, but they think that behavior is another story. I showed this to be wrong. When I found that either of two methods could cause an increase in sexual behavior 1) reducing authoritarianism, and thus enhancing creativity or 2. exposing subjects to a good sexual fantasy (Eisenman, 1982). This work builds on Grossman and Eisenman (1971), where we showed that reducing authoritarianism increased creativity, and increasing author- itarianism reduced creativity.

I found that a conservative, somewhat religiously based abstinence program could change school student's sexual attitudes in the direction of thinking they had more control of their sexual behavior than they previously believed, and should abstain (Eisenman, 1994) . But I did not know if only attitudes were changed, and not behavior. Only in Eisenman

(1982) was it shown that sexual behavior could be changed. One of my most important findings is that while first borns are more creative among males, the opposite is true among females. Later born females are more creative than first born females (eg. Boling, Boling & Eisenman, 1993). Many have realized that first borns are more likely to be creative than later-borns, but if you look at the findings, these usually only concern males.

Among females, the first born gets perhaps too much socialization, and that is harmful for producing a questioning somewhat rebellious person, who can come up with creative thoughts and productions. Too often, sex differences are not considered when research is conducted or reported.

The War on Drugs has been going on for some time. Years ago, we found that high schoolers who used marijuana tended to be more creative than nonusers (Victor, Grossman & Eisenman, 1973). So the putting of people in prison for using drugs may be putting a disproportionate number of creative people behind bars.

From working in a prison, I have some idea of what really happens, as opposed to what they want you to think goes on. I have written some articles critical of the prison organization (eg. Eisenman, 1991b, 1993a, 1995). I have also written about sex offenders and other types of criminals that I have encountered (Eisenman, 1994b, in press) I have shown that sex offenders can be divided into those who primarily use violence to get their way versus those who use what I call seduction (although there is violence implicit in what they do), and that the two groups speak differently in psychotherapy about victims, with the violent sex offenders more likely to denigrate victims (Eisenman,1993, 1997) . We need to pay more attention to the victim. In courts, the district attorney represents the state and the defendant has attorneys. No one, for the most part, represents the victim, and the focus of empathy is on the defendant. The victim has become forgotten. This is wrong.

I am interested in multicultural issues. My dad was a civil rights activist in Savannah, Georgia in the 1950's and 1960's working for the rights of African Americans. He was a self- appointed activist; he worked alone, organizing groups, meeting with ministers and politicians to get things done. As a result, I have been strongly concerned about equal rights for all. I described some of my dad's work and the effects—not all good—on our family (Eisenman, 1997). I developed a brief coaching method that helped African Americans improve their college exam scores (Eisenman, 1991c). I have been fascinated with how people accept the ideas of David Duke, a former head of the Ku Klux Klan, who now dresses in suits instead of Klan robes and says his earlier actions were youthful indiscretions. He seems like the same old racist to me, but often not to others (Eisenman, 1992b, 1993c, Eisenman, Girdner, Burroughs & Routman, 1993). Part of his appeal may be that he is one of the few people who consistently and publicly speak out against affirmative action and for the rights of whites, which many whites would like to do, but are afraid to do because of the likely negative consequences.

I have developed the concept of deviance, which is usually employed only by sociologists and generally meant only negative things. However, a few of us believe that deviance can refer to something positive, e.g. creativity. I wrote about this in my creativity/crime/ deviance book (Eisenman, 1991a).

Sometimes society tolerates too much deviance of the negative kind, and as a result, society is a worse place to live. However, sometimes deviance is created when little or nothing wrong has been done, as in the witch hunts. A modern day witch hunt often occurs when someone is charged with sexual harassment.

The term is so vague that anything can be said to be sexual harassment and the charged person is often almost as good as convicted, once accused. I presented three case histories of university faculty charged with sexual harassment (Eisenman, 1999). As far, as I know, I am the only person to interview, in detail, people charged with sexual harassment and to show the psychological harm to them, in terms of their experiencing anxiety and depression, to the possible point of considering suicide. They also face employment harm, as the sexual harassment charge often causes one to lose employment, and to be non-employable, as others react negatively to the stigma they now have. Yet the charges are often based on very flimsy evidence or involve things that many would not see as harmful at all

My research is often motivated by attempting to solve a mystery, or being angry about how something is presented, when I think it should be seen in an entirely different light. There is nothing wrong with an anger that results in more meaningful information being gathered, as opposed to an anger that cripples you from work or that leads to attacks on others.

REFERENCES

Bernard, J.L., & Eisenman, R. (1967). Verbal conditioning in sociopaths with social and monetary reinforcement. *Journal of Personality and Social Psychology. 6*, 303-306

Boling, S.E., Boling, J.L., & Eisenman, R. (1993). Creativity and birth order/sex differences in children. *Education, 114*, 224-226; 246.

Grossman, J.C. & Eisenman, R. (1971). Experimental manipulation of authoritarianism and its effects on creativity. *Journal of Consulting and Clinical Psychology, 36*, 328-344.

Eisenman, R. (1982). Sexual behavior as related to sex fantasies and experimental manipulation of authoritarianism and creativity. *Journal of Personality and Social Psychology, 43*, 853-860.

Eisenman, R. (1991a). *From crime to creativity : Psychological and social factors in deviance*. Dubuque, IA : Kendall/Hunt.

Eisenman, R. (1991b). I worked in a prison: An insider's story. *Psychology : A Journal of Human Behavior, 28*, (3/4), 22-26.

Eisenman, R. (1991c). Successful coaching of black students on college exam questions. *Education, 112*, 155-156.

Eisenman. R. (1992a). Creativity in prisoners: Conduct disorders and psychotics. *Creativity Research Journal, 5*, 175-181.

Eisenman, R. (1992b). Creativity, social and political attitudes and liking or disliking David Duke. *Bulletin of the Psychonomic Society, 30*, 19-22.

Eisenman, R. (1993a). Authoritarianism,, creativity, and other aspects of a prison organization. *Journal of Police and Criminal Psychology, 9*, (2), 77-79.

Eisenman, R. (1993b). Denigration of a victim in group psychotherapy by violent vs. seductive sex offenders. *Psychological Reports, 72*, 413-414.

Eisenman, R. (1993c). Student attitudes toward David Duke before and after seeing the film " Who is David Duke?" *Bulletin of the Psychonomic Society, 31*, 37-38.

Eisenman, R. (1994a). Conservative sexual values: Effects of an abstinence program on student attitudes. *Journal of Sex Education and Therapy, 20*, 75-78.

Eisenman. R. (1994b). Sexual addiction: A look at the concept and how well it explains sex offenders. *L.A.E. Journal of the American Criminal Justice Association, 55* (1 & 2), 25-29.

Eisenman, R. (1995). Unethical treatment of prisoners. *Ethics and Critical Thinking Journal, 34* (issue 404), 1-10.

Eisenman, R. (1997). Denigration of a victim in individual psychotherapy by violent vs. seductive child molesters. *Psychological Reports, 81*, 1276-1278.

Eisenman, R. (1999). Sexual harassment charges against university faculty: Three case histories. *Journal of Information Ethics, 8* (2) , 59- 75.

Eisenman, R. (in press). Why some men are sex offenders. *Social Justice Review.*

Eisenman, R., Girdner, E.J., Burroughs, R., & Routman, M. (1993). Attitudes of Mississippi College students towards David Duke before and after seeing the film " Who is David Duke" ? *Adolescence, 28*, 527-532.

Victor, H.R., Grossman, J.C. & Eisenman, R. (1973). Openness to experience and marijuana use in high school students. *Journal of Consulting and Clinical Psychology, 41*, 78-85.

Differential Processes of Emotion Space Over Time

Takuma Takehara and Naoto Suzuki
Doshisha University

This study investigated whether the structure of emotion space would change over time. Six sets of basic facial expressions were obtained from Ekman & Friesen's (1976) Pictures of facial affect. We divided each expression process into ten temporal segments. Facial images were morphed by a computer, each consisting of a series of faces differing by a constant physical amount, as they change from the neutral expression to the maximal. Participants rated those facial images by a 6-point Likert-type scale. Multidimensional scaling and cluster analysis revealed that various emotion categories fell into the same cluster at various steps from the segment 1 to 4, and that positions of emotion categories formed a circumplex at the segment 5, and that the circumplex pattern was fixed from the segment 6 to 10. It was suggested that there was a critical point of the change of emotion space between segment 4 and 5.

Facial expressions seem to offer the most readily available information about social interactions in our daily life. So far, many researchers have attempted to clarify various aspects about facial expressions. For example, Esteves and Ohman (1993) focused on the recognition thresholds for specific emotions (see also; Kirouac & Doré, 1983, 1984; McAndrew, 1986). Using the backwardly masked faces by neutral facial displays, they showed that confident recognitions of facial expressions required about 100-250 msec., with shorter time for happy than for angry expressions. Hess and Kleck (1994) mentioned on the discrimination between posed and spontaneous expressions (see also; Zuckerman, Hall, DeFrank, & Rosenthal, 1976). They discussed that a measure of the level of perceived honest demeanor of the stimulus persons based on their neutral expressions was found to relate to perceivers' accuracy in discriminating posed and spontaneous expressions. Ekman (1992, 1994) argued that emotions might be considered discrete in a sense and also emphasized the categorical perception of "basic emotions". Other researchers have arrived at the same conclusion (e.g., Ekman & Friesen, 1975; Ekman, Friesen, & Ellsworth, 1982; Izard, 1994; Tomkins, 1962, 1963). In addition, Ekman (1980) discussed the idea that there were some facial expressions of

---

Author info: Correspondence should be sent to: Takuma Takehara, Department of Psychology, Doshisha University, Kamigyo-ku, Kyoto, 602-8580, Japan. takehara@psychology.doshisha.ac.jp.

North American Journal of Psychology, 2001, Vol. 3, No. 2, 217-228.

© NAJP

emotion which were universal and Buck (1988) argued that facial expressions were correctly recognized in widely different cultures.

Furthermore, Russell and Bullock (1985, 1986) tried to represent dimensional structure of the face recognition from preschoolers to adults, and Takehara and Suzuki (1997) also showed the dimensionality of the face recognition using morphed facial expressions as stimuli. Takehara and Suzuki (1997) asked participants to rate the morphed facial images on a Likert-type scale, then they derived a two-dimensional solution. Similar to Russell and Bullock (1985, 1986), two axes of Pleasure-Displeasure and intensity of arousal were found. However, their result showed that participants tended to confuse fear expression with surprise expression. About confusion of emotions, Ekman, et al. (1982) also argued that fearful and surprised expressions were easily confused.

Although there have been many studies about facial expressions as described above, comparatively little attention has been directed to the question of whether other types of information might be communicated by facial expressions of emotions (Edwards, 1998). One of the important pieces of information that should not be ignored is the "time series" as Ekman (1984) mentioned that emotion had an onset, a peak, and an offset, suggesting the importance of the time series[1]. In this study, we define the temporal change from the neutral to the prototype as the time series. Edwards (1998) also argued importance of temporal cues embedded in facial displays. It is, however, true that many studies exploring emotion recognition or perception typically tend to address only to the final states of the expressive process or near the maximal expression and do not pay attention to the subtle and ambiguous change of facial information over time. Edwards (1998) noted the importance of temporal changes (i.e., time series) or cues of facial expressions and tested the hypothesis that individuals could detect extremely subtle dynamic cues that convey the temporal status of facial expressions of emotions. Edwards found the following interesting facts: (1) individuals could detect extremely subtle dynamic cues in a facial expression; (2) individuals were most sensitive to the temporal characteristics of the early stages of an expression. It is interesting to note that individuals are sensitive to the temporal characteristics of the early stages of an expression while the expressions at the early stage have less amount of emotional information.

As for the circumplex model of affect, it is considered that all six basic emotion categories might not have clear boundaries between them, even for the maximal (i.e., prototype) expressions (Takehara & Suzuki, 2000). This suggests that each emotion may have the other emotional components rather than only one. For example, the expression between Happiness and Surprise on the perimeter of a circumplex may have both emotional components of Happiness and Surprise (e.g., Takehara & Suzuki, 1997, 2000). The less emotional information a face has, the more likely the face would have other emotional components. As suggested by Takehara and Suzuki (1998) and Suzuki and Takehara (2000), research on the structure of emotion recognition (i.e., emotion space) over time (i.e., from an onset to a peak) may indicate that emotion boundaries at the beginning of expressions might be less distinguishable than at the peak. In other words, emotion space might have an ambiguous prototype at the beginning, and as time elapses, it may be differentiated into particular structural patterns.

Several studies have used spontaneous facial expressions as stimuli, and there seem to be a few fundamental problems associated with the procedure. First, it is difficult to standardize the duration of each emotion because each has different duration from an onset to the peak. For example, an expression of surprise has a shorter duration than that of sadness. Second, since each spontaneous expression seldom has the same orientation of such attributes as intensity, distinctness and temporal phase, it is difficult to prepare stimuli which are comparable on these attributes. This problem is more important than one may acknowledge because the stimuli need to be standardized and controlled for the comparability of outcomes. To mitigate this problem, we used the Ekman and Friesen (1976) face set and generated interpolated facial images from a neutral expression to its peak, using a morphing technique.

We hypothesize that if each emotion had several other emotion components, the emotion space would be complex and ambiguous at the beginning of the expression, and that it would be differentiated into a particular spatial pattern over time. Thus, the goal of this study is to investigate whether the emotion space will be differentiated from the ambiguous recognition structure to a more typical one over time. Regarding the task, we asked participants to rate the morphed facial images by a Likert-type scale. Then, based on the dissimilarities, the relationships of faces at each temporal segment were represented in the two-dimensional psychological space by multidimensional scaling (MDS) and cluster analysis.

METHOD

Stimuli

In order to generate standardized facial stimuli, we used the Ekman and Friesen (1976) series (Face JJ) because all facial stimuli have consistent quality, angle, and orientation, with standardized pose and controlled lightning. In the case of selecting emotion categories, we chose six basic emotions of Happiness, Sadness, Anger, Fear, Disgust, and Surprise, and one Neutral which were considered to cover a universal range of emotion[2]. To introduce time sequence, we used morphing, which can generate interpolated images between two parent images. We divided each expression process into ten temporal segments, and chose ten corresponding facial images from the beginning to the end (i.e., maximum) as stimuli. Relationships between four of the ten temporal segments and the corresponding morphed facial image are shown in Figure 1. All the resulting facial images, running from the neutral expression to the maximal prototype expression, provided continuity and the same photographic quality. In total, 60 facial images (6 basic emotion categories × 10 segments) were generated and all were used in this study. In addition, one geographic pattern image, which had no emotional information, but consisting of a triangle, a circle, and a square, was generated.

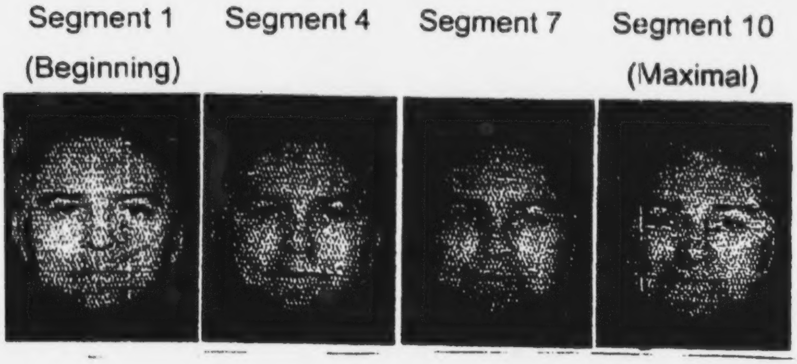


FIGURE 1 Facial images corresponding to four of the 10 temporal segments. This example shows the Happiness continuum.

Participants

The participants were 43 undergraduate students (13 males and 30 females ranging in age from 18 to 21 years with the mean of 18.8 years). They volunteered to be subjects in order to get a credit for a psychology course. All had normal or corrected-to-normal vision.

Procedure

Sixty generated facial images were displayed one at a time in a random order on a 14.1-inch LCD with 240×320 pixels. Note that an emotionless geographic pattern image was presented before each facial

BEST COPY AVAILABLE

image in order to minimize any influence of the previous facial image on the current image. The distance between a participant and a monitor was about 50 cm. Data were collected by a Compaq personal computer (DESKPRO; Pentium II Processor 350 MHz) with a two-button mouse and by a rating software made by an Inprise C++ compiler (Borland C++ Builder Version 4). Participants were asked to rate each facial image on the 6-point Likert-type scale (from "not at all" to "so very much") about each item of the six basic emotions of Happiness, Sadness, Anger, Fear, Disgust, and Surprise by clicking the mouse. By single-clicking of the "Next Image" button on the window, an emotionless geographic pattern image was presented for two seconds, followed by a next facial image. No feedback was given. Before the rating, two practice trials were given each subject.

RESULTS

MDS refers to a set of procedures used to depict proximities spatially between phenomena and detects hidden structure underlying complex cognitive constructs (Kruskal & Wish, 1978). This procedure provides a visual representation of the relations among the stimuli by placing them in a geometric space of specified dimensionality. One possible area of application for MDS is facial expressions of emotion and a number of stimulus sets have been studies using a wide range of procedures to collect the data and a variety of MDS techniques (Bimler & Kirkland, 1997). MDS has been frequently applied to studies on the structure of emotion (e.g., Abelson & Sermat, 1962 ; Katsikitis, 1997; Osgood, 1966; Russell, 1980; Russell & Bullock, 1985, 1986; Takehara & Suzuki, 1997). In the current study, rating data were used to calculate a matrix of dissimilarities among six emotion categories for each temporal segment, and MDS was carried out to find coordinates of the six emotion categories in the Euclidean space at each of ten temporal segments, respectively. The ALSCAL procedure from the Statistical Package for the Social Science (SPSS) was used, and extracted two solutions. Figure 2 shows its two-dimensional representation of the six emotion categories at the maximal expression. Greater dissimilarity between two facial images is represented by their greater distance. For the two-dimensional configuration, each stress and *R*-square values are shown in Table 1.

Kruskal and Wish (1978) and Coxon (1982) recommended the use of cluster analysis to aid the interpretation of the MDS solutions and we expected distinct patterns of relationships among the six emotion categories. Therefore, we employed hierarchical cluster analysis from SPSS, using the average linkage of squared distances between groups. We incorporated six variables in cluster analysis: Happiness, Sadness, Anger, Fear, Disgust, and Surprise. In each successive stage of hierarchical clustering, the distance between clusters was calculated and the closest clusters were combined. Thus, clusters combined at the first stage were closer to one another than those combined at the second stage. An example of a resultant cluster is presented in Figure 2. The circles, including deformed circles, depicted in Figure 2 represent the clustering schedule obtained from cluster analysis. The smaller inner circles represent the successive clustering in the hierarchical clustering scheme. Then, in order to see the detailed relations of emotions and changes of emotion space over time, let us look at each temporal segment.

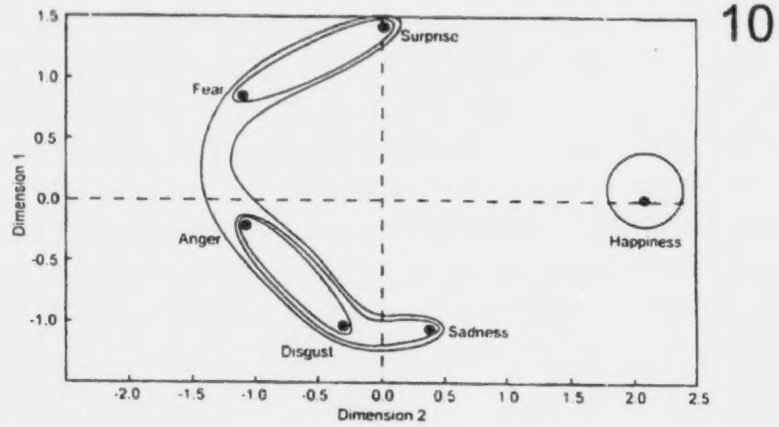


FIGURE 2. An example of a two dimensional solution at each of ten temporal segments. Clusters are presented at the same time. The number depicted on the upper-right of this graph represents the segment corresponding to Figure 1.

At Segment 1 (i.e., beginning of the expression), two-dimensional solutions and clusters were not distinguishable. "Sadness" and "Surprise" are combined first. Next, "Happiness," positive emotion, was combined with "Sadness-Surprise" cluster, and "Disgust" and "Fear" were combined into a single cluster. Finally, "Anger" was combined with "Sadness-Surprise-Happiness" cluster. Interpretation of both axes is difficult because they are mixed with negative and positive emotions.

At Segment 2, "Disgust" and "Surprise", and "Anger" and "Fear" were combined respectively. Then, "Sadness" was added to the "Disgust-Surprise" cluster, and finally emotions, except "Happiness," formed a single cluster. It is still difficult to interpret the axes.

At Segment 3, "Disgust" and "Surprise" were combined first. Next, "Fear" was added to the "Disgust-Surprise" cluster, and "Sadness" was added to the "Disgust-Surprise-Fear" cluster, and finally, "Happiness" was added to the "Disgust-Surprise-Fear-Sadness" cluster. Again, it is still difficult to interpret the axes.

TABLE 1 Stress and R-Square Values at Each of Ten Temporal Segments

Segment		Stress	R-Square
1	(beginning)	.063	.973
2		.020	.998
3		.079	.978
4		.079	.957
5		.142	.841
6		.114	.858
7		.169	.796
8		.157	.805
9		.198	.772
10	(maximum)	.186	.771

At Segment 4, "Disgust" and "Fear" were first combined. Next, "Sadness" was added to the "Disgust-Fear" cluster. Then, "Surprise" and "Happiness" were combined, and finally, "Anger" was combined with the "Disgust-Fear-Sadness" cluster. For this time segment, the horizontal axis may be considered to be the dimension of pleasure and displeasure, and the vertical axis intensity of arousal. Names of the axes, however, leave some uncertainty, because "Happiness" which is an absolutely positive emotion, for example, is combined with "Surprise" which may not contain any ingredients of positive emotion.

At Segment 5, "Disgust" and "Anger" were first combined, and then "Fear" and "Surprise" were combined. "Sadness" was added to the "Fear-Surprise" cluster, and finally, all emotions except "Happiness" fell into the single cluster. At this time segment of time, the two-dimensional solution shows approximately a circular order of emotions as Russell (1980) suggested. Moreover, only "Happiness" makes its own cluster. The results show that the horizontal axis can be considered the dimension of pleasure and displeasure, and that the vertical axis the intensity of arousal in terms of relationships of clusters and relative positions of emotion categories.

From Segment 6 to Segment 10 (i.e., the maximal expression), there were no changes to the circular order and clustering steps. In a nutshell, relative positions of emotion categories were similar to those of Russell's (1980)[3], and "Anger" was combined with "Disgust" in the first step, "Fear" was combined with "Surprise", "Sadness" was added to the

"Anger-Disgust" cluster, then all emotions except "Happiness" fell into a single cluster. As for the interpretation of both axes, the horizontal axis is considered the dimension of pleasure and displeasure, and the vertical axis the intensity of arousal, identical to those at Segment 5.

DISCUSSION

As Edwards (1998) pointed out, the notion that the temporal information might be an important factor of emotion recognition has never been examined systematically. As did Edwards, we considered "time" to be an important factor, and investigated how individuals' emotion space would change over time. For our experiment, we generated morphed facial images of each of six basic emotions running from the neutral expression to the maximal one, and we then divided the expression process into ten temporal segments, and chose ten corresponding facial images from the beginning (neutral) to the end (maximal). We only selected face JJ, because we wanted to eliminate the effect of or influence by other faces. This is not a major problem because face JJ is used in many studies (e.g., Takehara & Suzuki, 1997; Young, Rowland, Calder, Etcoff, Seth, & Perrett, 1997).

Let us discuss external validity of our artificial technology called morphing. Among all facial images, only the neutral and six prototypes were real images, meaning that most images used were artificially generated. Although those images do *not* exist in nature, the advantage of using morphed images is that it is easy to generate any interpolated images, making it possible to control the amount of physical changes. The best of all is that morphing provides us with a means of generating objective images. Recently, a number of researchers have used the procedure of morphing (e.g., Beale & Keil, 1995; Takehara & Suzuki, 1997, 2000; Tanaka et al., 1998; Young et al., 1997). While morphing technology may degrade the quality of the images, thus making emotions less discernable than actual images, this is clearly not a major problem, for the current findings show that emotion space became stable at Segment 5, indicating easy recognition of "generated" faces.

As for the handling of the time variable, the current study has shown successful grasp of the time as an important factor for the study of emotion. It would be true that the randomization of facial images can enhance control. However, its combination with the temporal factor can hardly be justified. Genuine temporal recognition of facial expressions appears to be context driven, making the preceding image vitally important for the following facial image. If face recognition was influenced by the context so much, however, the configuration of emotion space over time may never reach a distinctly stable stage. The current study did not use any context for emotion. But, the study successfully identified a stable configuration of emotion space, shortly after Segment 4 and definitely at Segment 5, which lasted until the last Segment. It is interesting to note that man can recognize the temporal changes of the face without any context information.

To reiterate, distance relationships among the six emotion categories were transformed to two-dimensional emotion space for each time segment by MDS and cluster analysis was then applied to the configuration to examine the differential processes of emotion space. This visual examination of the temporal processes has proven to be very useful.

It revealed that various emotion categories fell into various clusters from Segment 1 to Segment 4, that the positions of the six basic emotion categories formed a circumplex at Segment 5, and that this circumplex pattern continued stable from Segment 6 to Segment 10. The current findings agree with those of Takehara and Suzuki (1998) who reported a change of each emotion category from an ambiguous prototype at the beginning of the temporal process into a particular emotion category at the end. The fact that emotion space is completely differentiated between Segment 4 and Segment 5 (i.e., differentiated by the midpoint through the expression process) suggests that individuals are able to identify the other's facial expressions, which are at a comparatively early stages of the expression process. This might be consistent with Edwards (1998) who discussed that sensitivity to the temporal properties of an emotional expression may be greatest for the early stages of the expression.

With respect to the interpretation of axes in two-dimensional space, neither distance relationships among emotion categories by MDS nor classification relationships by cluster analysis can offer anything defensible before Segment 5. For example, because a positive emotion Happiness and a *non-positive* emotion Surprise fell into the same cluster at Segment 4, we could not interpret the axes in any meaningful way. It might be reasonable to say that both axes could not be interpreted for the early stages of the expression because each emotion category started with several other emotional components before they were gradually differentiated. In contrast, after Segment 5, the horizontal axis was considered the dimension of pleasure and displeasure, and the vertical axis the intensity of arousal.

Studies which focused only on the maximal expressions would never find that emotion categories get differentiated with other various emotional components as time elapses at the early stages of the expression. The current study demonstrated not only the importance of the research paradigm over time, but also the possible existence of a critical time point of emotion space between Segment 4 and Segment 5. The findings might be useful to study about micro-expressions of faces.

Although we divided the expression process into ten temporal segments, more segments may be needed to investigate the differential processes of emotion space over time if we were to use micro-expressions of faces.

REFERENCES

Abelson, R. P., & Sermat, V. (1962). Multidimensional scaling of facial expressions. *Journal of Experimental Psychology, 63*, 546-554.

Beale, J. M., & Keil, F. C. (1995). Categorical effects in the perception of faces. *Cognition, 57*, 217-239.

Bimler, D., & Kirkland, J. (1997). Multidimensional scaling of hierarchical sorting data applied to facial expressions. *Scandinavian Journal of Psychology, 38*, 349-357.

Buck, R. (1988). *Human motivation and emotion* (2nd. ed.). New York: Wiley.

Burton, A. M., Wilson, S., Cowan, M., & Bruce, V. (1999). Face recognition in poor-quality video: Evidence from security surveillance. *Psychological Science, 10*, 243-248.

Coxon, A. P. M. (1982). *The user's guide to multidimensional scaling*. London: Heinemann Educational Books.

Edwards, K. (1998). The face of time: Temporal cues in facial expressions of emotion. *Psychological Science, 9*, 270-276.

Ekman, P. (1980). *The face of man: Expressions of universal emotions in a New Guinea village*. New York: Garland STPM Press.

Ekman, P. (1984). Expression and the nature of emotion. In K. R. Scherer & P. Ekman (Eds.), *Approaches to emotion* (pp. 319-343). Hillscale, NJ: Erlbaum.

Ekman, P. (1992). Are there basic emotions? *Psychological Review, 99*, 550-553.

Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. *Psychological Bulletin, 115*, 268-287.

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face*. Englewood Cliffs, NJ: Prentice-Hall.

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., Friesen, W. V., & Ellsworth, P. (1982). *Emotion in the human face. (2nd ed.)* Cambridge, MA: Cambridge University Press.

Esteves, F., & Ohman, A. (1993). Masking the face: Recognition of emotional facial expressions as a function of the parameters of backward masking. *Scandinavian Journal of Psychology, 34*, 1-18.

Hess, U., & Kleck, R. E. (1994). The cues decoders use in attempting to differentiate emotion-elicited and posed facial expressions. *European Journal of Social Psychology, 24*, 367-381.

Izard, C. E. (1994). Innate and universal facial expressions: Evidence from developmental and cross-cultural research. *Psychological Bulletin, 115*, 288-299.

Katsikitis, M. (1997). The classification of facial expressions of emotion: A multidimensional-scaling approach. *Perception, 26*, 613-626.

Kirouac, G., & Doré, F. Y. (1983). Accuracy and latency of judgment of facial expressions of emotions. *Perceptual and Motor Skills, 57*, 683-686.

Kirouac, G., & Doré, F. Y. (1984). Judgment of facial expressions of emotion as a function of exposure time. *Perceptual and Motor Skills, 59*, 147-150.

Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling.* Newbury Park, CA: Sage.

McAndrew, F. T. (1986). A cross-cultural study of recognition thresholds for facial expressions of emotion. *Journal of Cross-Cultural Psychology, 17*, 211-224.

Osgood, C. E. (1966). Dimensionality of the semantic space for communication via facial expressions. *Scandinavian Journal of Psychology, 7*, 1-30.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161-1178.

Russell, J. A., & Bullock, M. (1985). Multidimensional scaling of emotional facial expressions: Similarity from preschoolers to adults. *Journal of Personality and Social Psychology, 48*, 1290-1298.

Russell, J. A., & Bullock, M. (1986). On the dimensions preschoolers use to interpret facial expressions of emotion. *Developmental Psychology, 22*, 97-102.

Suzuki, N., & Takehara, T. (2000, August). *Differential process of emotions II: A case of professional actors.* Poster session presented at the 11th conference of the International Society for Research on Emotions, Quebec City, Canada.

Takehara, T., & Suzuki, N. (1997). Morphed images of basic emotional expressions: Ratings on Russell's bipolar field. *Perceptual and Motor Skills, 85*, 1003-1010.

Takehara, T., & Suzuki, N. (1998). Differential processes of emotions. In A. Fischer (Ed.), *Proceedings of the 10th conference of the International Society for Research on Emotions* (pp. 328-331).

Takehara, T., & Suzuki, N. (2000). *Coexistence of Categorical and Dimensional Accounts for the Structure of Face Recognition: Representation of Psychological Phase Space.* Manuscript submitted for publication.

Tanaka, J., Giles, M., Kremen, S., & Simon, V. (1998). Mapping attractor fields in face space: The atypicality bias in face recognition. *Cognition, 68*, 199-220.

Tomkins, S. S. (1962). *Affect, imagery, consciousness: Vol. 1. The positive affects.* New York: Springer Verlag.

Tomkins, S. S. (1963). *Affect, imagery, consciousness: Vol 2. The negative affects.* New York: Springer Verlag.

Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., & Perrett, D. I. (1997). Facial expression megamix: Tests of dimensional and category accounts of emotion recognition. *Cognition, 63*, 271-313.

Zuckerman, M., Hall, J. A., DeFrank, R. S., & Rosenthal, R. (1976). Encoding and decoding of spontaneous and posed facial expressions. *Journal of Personality and Social Psychology, 34*, 966-977.

Author Note: We are grateful to Professor Paul Ekman for permission to use photographs from the Ekman and Friesen (1976) Pictures of Facial Affect series.

We thank Professor Shizuhiko Nishisato, University of Toronto, for checking of the draft article and making useful comments, Mary Katsikitis, Daniel Messinger, Michael Lyons, Monique Debonis, and Harald Wallbott for comments at the 11th ISRE conference. Thanks are also due to Professor Fumio Ochiai, Tezukayama University, for permission to collect data.

Footnotes

[1] There are a few studies using videotaped stimuli (e.g., Burton, Wilson, Cowan, & Bruce, 1999), but few studies have explored emotion identification or emotion concept.

[2] Selected identifiers of Face JJ were Happiness JJ-4-8, Sadness JJ-5-5, Anger JJ-3-12, Fear JJ-5-13, Disgust JJ-3-20, Surprise JJ-4-13, and Neutral JJ-3-4.

[3] The same circular order about emotion space was found in many studies (e.g., Takehara & Suzuki, 1997, 1998, 2000).

Healthy Perfectionism and Positive Expectations About Counseling

J. M. Oliver Brett A. Hart

Michael J. Ross Barry M. Katz[1]
Saint Louis University

This study examined the relation between perfectionism as measured by the Multi-Dimensional Perfectionism Scale and expectations about counseling (EAC's) as measured by the Expectations About Counseling-Brief Form in 200 undergraduates. Preliminary analyses indicated that Self-Oriented, Other-Oriented, and Socially-Prescribed Perfectionism were linearly related to all 17 facets of EAC's and that subscales of the EAC were the appropriate unit of analysis. Canonical correlation demonstrated a significant relation between dimensions of perfectionism and EAC's such that adaptive perfectionism appeared to be associated with positive EAC's for the client, the counselor, and counseling outcome and process. Results provide further support for a healthy type of perfectionism and suggest that it may indicate a good candidate for counseling.

The purpose of this study was to examine the relation between perfectionism and expectations about counseling in a large group of unselected undergraduate students. Several recent reports have suggested the existence of an adaptive as distinct from a maladaptive type of perfectionism (Hamacheck, 1978; LoCicero & Ashby, 2000; Parker, 1997; Rice, Ashby, & Slaney, 1998), and previous research suggests that positive expectations about counseling may be associated with better therapy process and outcome (Al-Darmaki & Kivlighan, 1993; Tinsley, Tokar, & Helwig, 1994; Tokar, Hardin, Adams, & Brandel, 1996). However, it remained to be discovered whether adaptive perfectionism could be detected in another undergraduate sample and, if so, whether such perfectionism might be related to expectations about counseling.

Current investigators seem to construe adaptive perfectionists as a subgroup of perfectionists (LoCicero & Ashby, 2000; Parker, 1997). The identification of adaptive perfectionism refines previous study of perfectionism, which had established that overall, perfectionism is a risk factor for maladjustment and psychopathology of many types. Current

---

Author info: Correspondence should be sent to: Dr. J. M. Oliver, Dept. of Psychology, St. Louis University, 221 N. Grand Blvd., St. Louis, MO 63103.

North American Journal of Psychology, 2001, Vol. 3, No. 2, 229-242.
© NAJP

interest in adaptive perfectionism establishes another link between contemporary research on perfectionism and previous early theorizing, which had hypothesized that perfectionism could be of two forms, either healthy or neurotic (Adler, 1956; Pacht, 1984).

Several recent empirical studies have detected adaptive perfectionists (LoCicero & Ashby, 2000; Parker, 1997; Rice, Ashby, & Slaney, 1998). Adaptive perfectionists have been found among college undergraduates (LoCicero & Ashby, 2000; Rice, et al., 1998) and gifted presecondary students (Parker, 1997). Adaptive perfectionists appear to set higher standards for academic performance and to have with higher levels of perceived self-efficacy, both general self-efficacy and social self-efficacy; however, they appear to be less concerned than others about discrepancies between their standards and their performance (LoCicero & Ashby, 2000). Adaptive perfectionists acknowledge higher levels of conscientiousness, but they also describe themselves in generally positive terms, such as agreeable and extraverted (Parker, 1997).

Perfectionism generally has been found to be associated with a wide range of maladjustment and psychopathology, including depression (Flett et al., 1991; Hewitt & Flett, 1990, 1991a, 1991b); suicidality (Hewitt, Flett, & Turnbull-Donovan, 1992); anxiety (Flett, Hewitt, & Dyck, 1989; Hewitt & Flett, 1990); and alcoholism, anorexia nervosa, bulimia, personality disorders, chronic pain, and obsessive-compulsive disorder (Hewitt & Flett, 1991a). Maladaptive perfectionists have likewise been found among typical college undergraduates (LoCicero & Ashby, 2000); college undergraduates with learning disabilities (Saddler & Buckland, 1995); medical and dental students (Ey, Henning, & Shaw, 2000); and gifted pre-secondary students (Parker, 1997). Maladaptive perfectionists appear to experience significant others as setting high standards for them, having high expectations of them, and being critical of them. They also tend to be concerned about discrepancies between standards and their performance (LoCicero & Ashby, 2000); to have decreased self-esteem (Rice, et al., 1998); to have lower perceived self-efficacy (LoCicero & Ashby, 2000); to procrastinate more and to be more depressed (Saddler & Buckland, 1995; Saddler & Sacks, 1993); and not to seek formal sources of help despite high levels of distress and to have more negative attitudes about mental health services (Ey, et al., 2000).

Hewitt and Flett (1991) have identified three specific and independent components of perfectionism that may be characterized by their origin (the self or others) and their target (the self or others). Self-oriented perfectionism is perfectionism that originates in the self and is directed toward the self; it entails setting high standards for the self, striving to meet these standards, and stringent self-evaluations of one's own performance. Other-oriented perfectionism originates in the self but is directed toward others; it entails setting unrealistically high standards for significant others, placing importance on these others' being perfect, and stringently evaluating others' performance. Socially-prescribed perfectionism is perceived as originating in significant others and being directed toward the self; it entails perceptions that significant others set unrealistically high standards for oneself, pressure one to be perfect, and evaluate one's own performance stringently. As indicated above, recent investigators have found adaptive perfectionists to be characterized by self-oriented perfectionism, in contrast to maladaptive perfectionists, who are characterized by socially-prescribed perfectionism.

Expectations of counseling constitute one indicator of potential clients' willingness to seek help and of their likely success in therapy. Individuals who harbor more positive expectations of counseling are more willing to seek help and, once in therapy, are more likely to get off to a good start. Recent research suggests that favorable expectations about counseling have a positive impact on the course and outcome of therapy. For example, clients with higher expectations are more involved in the first interview (Tinsley et al., 1994) and have a stronger working alliance (Al-Darmaki & Kivlighan, 1993; Tokar et al., 1996).

Many measures have been developed to measure expectations about therapy and counseling (Hayes & Tinsley, 1989). The Expectations About Counseling-Brief Form (EAC-B; Tinsley, 1982) is the measure that simultaneously is the most widely-used and whose psychometric properties are best known. Tinsley and his colleagues (Tinsley, Workman, & Kass, 1980) designed the parent instrument, the longer *Expectations About Counseling-Full Form*, rationally to represent expectations about counseling in three distinct realms: 1) client attributes; 2) counselor attributes; and 3) counseling process and outcome. Tinsley and others (Hortman, 1984; Tinsley et al., 1980) factor analyzed the EAC to discern the empirical structure as distinct from the theoretical structure of the scale and extracted four dimensions: 1) Personal Commitment (investment of the client in the counseling process and outcome); 2) Facilitative Conditions (aspects of the counseling relationship and process that foster progress); 3) Counselor Expertise (attitudes and behaviors of the counselor that foster progress); and 4) Nurturance (characteristics of the counselor that are experienced as supportive by the client). The brief form of the instrument retained the factor structure and psychometric properties of the original form while decreasing participant time. The EAC-B has lent itself to application in empirical research concerned with counseling initiation, process, and outcome.

It seems plausible that individuals who harbor adaptive perfectionism may have a sense of personal control over initiating therapy and over its course and outcome; have positive, high standards for course and outcome; and have a sense of self-efficacy in bringing about a positive course and outcome, leading to overall positive expectations of counseling. It likewise seems plausible that individuals who harbor maladaptive perfectionism may lack a lack of a sense of control over initiating therapy; may sense that significant others or the therapist will set exacting standards for the course and outcome of therapy; and lack a sense of self-confidence in bringing about a positive course and outcome, culminating in overall negative expectations of counseling. Since undergraduate students as a group are selected for psychological health and high achievement strivings, we predicted that we would identify a type of adaptive, healthy perfectionism, and that it would be associated with positive expectations about counseling.

METHOD

Participants

The investigation was conducted in 200 undergraduate students at a mid-size private Catholic university in the Midwest. Generally, participants were about 70% female; between the ages of 18 and 21; about 80% Caucasian with substantial minorities of African-American and Asian students but little representation of other ethnic minorities; about 60% Catholic; 94% native-English-speaking; about 50% first year students with the remainder about equally divided among the other year levels. They represented a broad range of majors and family social class backgrounds as indicated by parental income.

Measures

Multi-Dimensional Perfectionism Scale (MPS): The MPS (Hewitt & Flett, 1989) is a 45-item self-report measure. Hewitt and Flett (1989) constructed the MPS theoretically to measure three aspects of perfectionism construed as relatively independent dimensions: self-oriented perfectionism; other-oriented perfectionism; and socially-prescribed perfectionism. Self-oriented perfectionism pertains to standards for the self; is postulated to constitute vulnerability to self-criticism, self-blame, and anxiety; and is assessed by items asking whether the respondent strives to be the best at everything he or she does. Other-oriented perfectionism pertains to standards for others; is postulated to predispose toward criticism and blame of others and feelings of anger; and is assessed by items asking whether the respondent cannot stand to see people close to him or her make mistakes. Socially-prescribed perfectionism pertains to the perception that others, typically significant others, have high and unreasonable standards for the self. It is postulated to predispose one toward a sense of constant pressure and anxiety; and is assessed by items asking whether the respondent perceives others as expecting more of him or her than the individual is capable of giving. Each sub-scale comprises 15 items.

Participants indicate their responses to each item on a 7-point Likert scale whose anchors are "strongly disagree" and "strongly agree." Scores on both the sub-scales and the total scale are expressed as the total of the items included, resulting in a range of 15-105 for each sub-scale and 45-315 for the total score.

Coefficient α's of .88, .74, and .81 have been reported for the self-oriented, other-oriented, and socially-prescribed perfectionism sub-scales, respectively (Hewitt & Flett, 1991a). Test-retest reliabilities across a three-month interval of .88, .85, and .75 have been reported for the self-oriented, other-oriented, and socially-prescribed perfectionism sub-scales, respectively (Hewitt & Flett, 1991a), indicating that perfectionism is quite a stable dimension of personality.

Construct validity of the MPS has been supported by principal-components factor analysis, which yielded three factors corresponding to the three dimensions of perfectionism proposed on a theoretical basis (Hewitt & Flett, 1991a). Convergent and discriminant validity of the MPS have been demonstrated in differential associations between the three dimensions of perfectionism and measures pertaining to standards for the self, for others, and for social situations, respectively.

Expectations About Counseling Scale-Brief Form (EAC-B): The EAC-B (Tinsley, 1982) was developed as a brief version of the original, 147-item self-report instrument. Correspondence between the brief and parent instruments is high as indicated by correlations between corresponding sub-scales and factors which range from .81 to .95. The brief form contains 66 items making up 17 sub-scales that contribute to factors and one experimental subscale. Of the 17 sub-scales of the EAC-B, three measure expectations about client attributes (Motivation, Openness, and Responsibility); 11 measure expectations about counselor attributes (Acceptance, Confrontation, Directiveness, Empathy, Genuineness, Nurturance, Self-Disclosure, Attractiveness, Expertise, Tolerance, and Trustworthiness); and three measure expectations about process and outcome (Concreteness, Immediacy, and Outcome).

Participants respond to each item on a seven-point Likert scale whose anchors are "not true" and "very true". Scores are reported as averages of the items making up each subscale and each factor of the instrument, yielding a range for both subscales and factor scores ranging from 0 to 7.

Coefficients α's for the 17 scales have been reported to range from .69 to .82 with a median value of .76 (Tinsley, 1982). Test-retest reliabilities of the sub-scales across a two-month interval have been reported to range from .47 to .87 with a median value of .71 (Tinsley, 1982).

Construct validity of the EAC-B has been supported in an investigation using a thought-listing technique to tap participants' spontaneous cognitions elicited by each item (Tinsley & Westcot, 1990). Discriminant validity has been demonstrated in a factor-analytic investigation of seven instruments comprising 33 subscales and 23 dimensions of counseling, from which two sets of factors were extracted: three factors pertaining to expectations about counseling and three factors pertaining to perceptions of counseling (Hayes & Tinsley, 1989). Further, the EAC-B has been reported to diverge from measures of student developmental task, psychological problems, and counseling readiness (Tinsley, Hinson, Holt, & Tinsley, 1990, 1992).

Demographic Inventory: Demographic variables surveyed included gender; year in school; major; ethnicity; religious affiliation; enrollment status (full-time vs. part-time); employment status; residence; estimated parental income; and cumulative GPA.

Procedure

Participants were recruited through undergraduate college courses. Participants were promised complete anonymity and completion of the composite questionnaire was taken to indicate informed consent. Participants were offered extra credit in their course as an incentive for their participation.

RESULTS

Means, standard deviations, and coefficient α's of psychological variables are displayed in Table 1. Means and standard deviations on the MPS were similar to those reported for a community sample of 100 males and 99 females (Hewitt, Flett, Turnbull-Donovan, & Mikail, 1991); means and standard deviations on the subscales of the EAC-B closely resembled those reported in 40 undergraduates reared in the United States (Yuen & Tinsley, 1981). Coefficient α's generally indicate high internal consistency of each sub-scale.

TABLE 1 Means, Standard Deviations, and Coefficient Alphas of Psychological Variables

Psychological Variable	*M*	*SD*	α
Multi-Dimensional Perfectionism Scale			
Self-Oriented Perfectionism	74.30	15.33	.90
Other-Oriented Perfectionism	59.39	9.56	.69
Socially-Prescribed Perfectionism	53.39	12.81	.83
Total Score	187.11	28.81	- -
Expectations About Counseling-Brief Form			
Motivation	4.70	1.32	.80
Openness	5.29	1.25	.80
Responsibility	5.76	0.92	.70
Acceptance	5.05	1.28	.74
Confrontation	5.32	1.12	.83
Directiveness	4.23	1.38	.71
Empathy	3.57	1.44	.76
Genuineness	6.25	0.89	.73
Nurturance	5.78	1.01	.73
Self-Disclosure	3.49	1.53	.77
Attractiveness	4.15	1.26	.82
Expertise	5.48	1.12	.73
Tolerance	4.96	1.21	.68
Trustworthiness	5.91	1.07	.82
Concreteness	5.60	1.01	.82
Immediacy	5.24	1.04	.72
Outcome	5.52	1.01	.75

Note: *N*'s = 197-200.

Preliminary analyses were conducted to ascertain whether relations between perfectionism and expectations about counseling were linear or curvilinear. It could be argued that perfectionism would be curvilinearly related to expectations about counseling and that the people who are moderately perfectionistic would have the highest expectations for counseling, whereas those who were least perfectionistic and most perfectionistic would have relatively lower expectations for counseling. The shape of relations between dimensions of perfectionism and components of the EAC-B was examined by means of one-way analyses of variance. Each dimension of perfectionism was taken in turn as the independent variable after trichotomizing the dimension using cut-points one standard deviation above and one standard deviation below the mean; sub-scales and summarizing components of the EAC and maladjustment were analyzed as dependent variables. In each case in which the relation between a dimension of perfectionism and a facet of the EAC was significant, the relation as linear: higher levels of perfectionism were associated with higher expectations on the EAC.

TABLE 2 Summary of Canonical Correlation Analysis of Dimensions of Perfectionism and Subscales of Expectations About Counseling

Variables	Structure Coefficient
Dependent Variables	
Responsibility	.52
Acceptance	.65
Confrontation	.30
Directiveness	.03
Genuineness	.62
Expertise	.42
Tolerance	.45
Trustworthiness	.77
Immediacy	.65
Outcome	.54
Independent Variables	
Self-Oriented Perfectionism	.65
Other-Oriented Perfectionism	.08
Socially-Prescribed Perfectionism	-.43

Factor analysis was used to determine the most appropriate way of summarizing and combining the 17 subscales of the EAC. Maximum-likelihood factor analysis specifying four factors and followed by both Varimax and Oblimin rotations was conducted to try to recover the four factors previously extracted by Tinsley and others (Hortman, 1984; Tinsley et al., 1980). Although a four-factor solution seemed indicated by both eigenvalues and a scree plot, the four factors taken together accounted for only about 57% of the variance, so the solution did not summarize the instrument particularly well. Of particular concern was the fact that the four factors did not correspond well to Personal Commitment, Facilitative Conditions, Counselor Expertise, and Nurturance, respectively: according to Harmon's coefficient of congruence; correspondences were .56, .36, .51, and .36 for the four pairs of factors, respectively. Further, the factors we extracted seemed to represent amalgams of client attributes, counselor attributes, and process and outcome; the factors seemed to lack psychological coherence.

For this reason, we turned to the 17 subscales of *Expectations About Counseling* as the appropriate unit of analysis. Coefficient α's of the subscales were generally high; M = .76, range = .68-.83.

Zero-order correlations with one-tailed tests of significance were calculated between the three dimensions of perfectionism and the 17 subscales of Expectations About Counseling. Of the 51 correlations, 12 were significant at or beyond the .05 level. Self-Oriented Perfectionism was significantly correlated with one client attribute (Responsibility [r = .16, p ≐ .013]; five counselor attributes (Acceptance [r = .26, p = .000]; Confrontation [r =.18, p = .005]; Expertise [r = .12, p = .044]; Tolerance [r = .24 p = .000]; and Trustworthiness [r = .21, p = .001]; and two characteristics of process/outcome (Immediacy [r = .16, p = .014]; and Outcome [r = .14, p = .024]). Other-Oriented Perfectionism was significantly correlated with two counselor attributes: Acceptance (r = .16, p = .011) and Tolerance (r = .14, p = .027). Socially-Prescribed Perfectionism was also significantly correlated with two counselor attributes: Directiveness (r = .12, p = .044) and Genuineness (r = -.13, p = .033). *Df*'s ranged from 193-196.

Canonical correlation analysis was conducted to examine the relations among all dimensions of perfectionism and the 10 subscales of the EAC-B that were significantly associated with at least one dimension of perfectionism. Results indicated that one of three possible canonical variates was significant at or beyond the .05 level of significance; Wilks λ= .76, df = 30, p = .008; however, according to the Stuart-Love Redundancy Index, this canonical variate explained only 3.5% of the variance in the dependent variables. F tests indicated that Self-Oriented and Socially-Prescribed Perfectionism were significantly related to subscales of the EAC-B and that of the 10 subscales of the EAC-B, all were significantly related to dimensions of perfectionism except Directiveness.

Structure coefficients of the canonical variate relating dimensions of perfectionism to facets of EAC's are displayed in Table 2. As shown in Table 2, Trustworthiness (an expectation about counselors) was most strongly and highly related to the canonical variate; Acceptance (an expectation about counselors), Immediacy and Outcome (characteristics expected of process and outcome) and Responsibility (an expectation about participants as clients) were also moderately related to the canonical variate. Genuineness, Expertise, Tolerance, and Confrontation (all expectations about counselors) were significantly related to the canonical variate, but to a lesser degree. Because this configuration of subscales of Expectations About Counseling spanned all three rationally-based components of expectations (client attributes, counselor attributes, and process and outcome), it seemed reasonable to term it simply Positive Expectations for Counseling. As shown in Table 2, Self-Oriented Perfectionism (describing standards set by and for the self) was positively and moderately related to the variate. Socially-Prescribed Perfectionism (describing feelings of pressure from demands placed on one by significant others) was moderately but negatively related to the variate; and Other-Oriented Perfectionism (describing expectations that others should be perfect) was unrelated to the variate. Because this configuration of dimensions so closely reflected previous characteristics of adaptive perfectionism, it likewise seemed reasonable to simply term the configuration Healthy Perfectionism. To summarize, canonical correlation analysis found a significant relation between participants' Healthy Perfectionism and Positive Expectations for Counseling.

DISCUSSION

In summary, in this college student sample, the four-factor structure that characterized expectations about counseling in two separate investigations did not emerge; although four factors could be extracted, they lacked psychological coherence. The fact that the four-factor structure could not be replicated together with generally high internal consistency of the individual sub-scales implied that individual sub-scales were the best units of analysis in this sample. Canonical correlation analysis demonstrated that dimensions of perfectionism were significantly related to facets of expectations. Perfectionism and expectations about counseling cohered in a pattern that suggest that adaptive perfectionism, or healthy positive striving, is significantly associated with positive expectations for all facets of counseling, including expectations of the client, the counselor, and the process and outcome of counseling.

To some extent our results might be seen as the "mirror image" of those of Ey et al. (2000), who found that distressed medical and dental students who failed to seek treatment had higher levels of socially-prescribed perfectionism and more negative attitudes toward mental health services. By contrast, we found that healthy perfectionism, characterized by a large component of self-oriented perfectionism and an absence of socially-prescribed perfectionism, was associated with positive expectations about counseling. By extension, our results suggest that the inverse of healthy perfectionism, characterized by a large component of socially-prescribed perfectionism and an absence of self-oriented perfectionism, might be associated with negative expectations about counseling and that these negative expectations might, in turn, discourage students from seeking counseling.

Our findings need to be qualified by the possibility that we may not have sampled students with extremely high or extremely low levels of perfectionism. Curvilinear associations between perfectionism and expectations about counseling may occur in samples with more extreme scores on either dimensions of perfectionism or facets of expectations of counseling. For example, it may be that some individuals who are extraordinarily perfectionistic have lower expectations of counseling; a sample of clients actually beginning counseling may include individuals who are more perfectionistic and harbor lower expectations.

It would be important to extend this investigation to samples of clients currently beginning counseling and/or on waiting lists at treatment centers. Such research might extend the range of perfectionism sampled and might help to clarify the nature of the relations among perfectionism and expectations about counseling at the upper ends of the ranges of perfectionism and at both ends of expectations about counseling. Further, as suggested by Ey et al. (2000), help-seeking may in itself be related to either of these variables.

Our results obtained in a college sample suggest that among university undergraduates, who constitute a group generally selected for psychological health and high motivation to achieve, there exists a constructive, adaptive type of perfectionism, and that this healthy perfectionism is associated with positive expectations for counseling. Our results join other recent reports of students who harbor adaptive perfectionism that has positive personality correlates. These results suggest that further research on perfectionism may find it productive to differentiate between adaptive or healthy perfectionism and maladaptive or neurotic perfectionism. They likewise suggest, following Ey et al. (2000), that it may be important for college personnel to understand that college students who harbor maladaptive or neurotic perfectionism may need the encouragement of psycho-education. They may also require positive role modeling, and/or normalizing of treatment to overcome their resistance to entering treatment if they become distressed.

REFERENCES

Adler, A. (1956). The neurotic disposition. In H. L. Ansbacher and R.R. Ansbacher (Eds.), *The Individual Psychology of Alfred Adler*. New York: Harper.

Al-Darmaki, F., & Kivlighan, D. M. (1993). Congruence in client-counselor expectations for relationship and the working alliance. *Journal of Counseling Psychology, 40*, 379-384.

Ey, S., Henning, K. R., & Shaw, D. L. (2000). Attitudes and factors related to seeking mental health treatment among medical and dental students. *Journal of College Student Psychotherapy, 14*, 23-39.

Flett, G. L., Hewitt, P. L., Blankstein, K. R., & Koledin, S. (1991). Dimensions of perfectionism and irrational thinking. *Journal of Rational-Emotive and Cogntive Behavior Therapy, 9*, 185-201.

Flett, G. L., Hewitt, P. L., & Dyck, D. G. (1989). Self-oriented perfectionism, neuroticism and anxiety. *Personality and Individual Differences, 10*, 731-735.

Hamacheck, D. E. (1978). Psychodynamics of normal and neurotic perfectionism. *Psychology, 15*, 27-33.

Hayes, T. J., & Tinsley, H. E. (1989). Identification of the latent dimensions of instruments that measure perceptions of and expectations about counseling. *Journal of Counseling Psychology, 36*, 492-500.

Hewitt, P. L., & Flett, G. L. (1989). The *Multidimensional Perfectionism Scale*: Development and validation. *Canadian Psychology, 30*, 339 (Abstract)

Hewitt, P. L., & Flett, G. L. (1990). Perfectionism and depression: A multidimensional analysis. *Journal of Social Behavior and Personality, 5*, 423-438.

Hewitt, P. L., & Flett, G. L. (1991a). Perfectionism in the self and social contexts: Conceptualization, assessment, and association with psychopathology. *Journal of Personality and Social Psychology, 60*, 456-470.

Hewitt, P. L., & Flett, G. L. (1991b). Dimensions of perfectionism in unipolar depression. *Journal of Abnormal Psychology, 100*, 98-101.

Hewitt, P. L., Flett, G. L., & Turnbull-Donovan, W. (1992). Perfectionism and suicide potential. *British Journal of Clinical Psychology, 31*, 181-190.

Hewitt, P. L., Mittelstaedt, W., & Wollert, R. (1989). Validation of a measure of perfectionism. *Journal of Personality Assessment, 53*, 133-144.

Hortman, J. W. (1984). A construct validation study of the Expectations About Counseling Questionnaire. (Doctoral dissertation, University of Alabama, 1984). *Dissertation Abstracts International, 44*(10), 2981A.

LoCicero, K. A., & Ashby, J. S. (2000). Multidimensional perfectionism and self-reported self-efficacy in college students. *Journal of College Student Psychotherapy, 15*, 47-56.

Pacht, A. R. (1984). Reflections on perfection. *American Psychologist, 39*, 386-390.

Parker, W. D. (1997). An empirical typology of perfectionism in academically talented children. *American Educational Research Journal, 34*, 545-562.

Rice, K., Ashby, J., & Slaney, R. (1998). Self-esteem as a mediator between perfectionism and depression: A structural equation analysis. *Journal of Counseling Psychology, 45*, 304-314.

Saddler, C. D., & Buckland, Jr., R. L. (1995). The *Multidimensional Perfectionism Scale*: Correlations with depression in college students with learning disabilities. *Psychological Reports, 77*, 483-490.

Saddler, C. D., & Sacks, L. A. (1993). Multidimensional perfectionism and academic procrastination: Relationships with depression in university students. *Psychological Reports, 73*, 863-871.

Tinsley, H. E. A. (1982). *Expectations About Counseling (copyright).*Unpublished test manual, Southern Illinois University at Carbondale, Department of Psychology, Carbondale, IL 62901.

Tinsley, H. E., Bowman, S. L., & Barich, A. W. (1993). Counseling psychologists' perceptions of the occurrence and effects of unrealistic expectations about counseling and psychotherapy among their clients. *Journal of Counseling Psychology, 40*, 46-52.

Tinsley, D. J., Hinson, J. A., Holt, M. S., & Tinsley, H. E. (1990). Level of psychosocial development, perceived level of psychological difficulty, counseling readiness, and expectations about counseling: Examination of group differences. *Journal of Counseling Psychology, 37*, 143-148.

Tinsley, D. J., Hinson, J. A., Holt, M. S., & Tinsley, H. E. (1992). A construct validation study of the *Expectations about Counseling-Brief Form*: Factorial validity. *Measurement and Evaluation in Counseling and Development, 24*, 101-110.

Tinsley, H. E. A., Tokar, D. M., & Helwig, S. E. (1994). Client expectations about counseling and involvement during career counseling. *Career Development Quarterly, 42*, 326-336.

Tinsley, H. E., & Westcot, A. M. (1990). Analysis of the cognitions stimulated by the items on the *Expectations About Counseling-Brief Form*: An analysis of construct validity. *Journal of Counseling Psychology, 37*, 223-226.

Tinsley, H. E. A., Workman, K. R., & Kass, R. (1980). Factor analysis of the domain of client expectancies about counseling. *Journal of Counseling Psychology, 27*, 561-570.

Tokar, D. M., Hardin, S. I., Adams, E. M., & Brandel, I. W. (1996). Clients' expectations about counseling and perceptions of the working alliance. *Journal of College Student Psychotherapy, 11*, 9-26.

Yuen, R. K. & Tinsley, H. E. A. (1981). International and American Students' Expectancies About Counseling. *Journal of Counseling Psychology, 28*, 66-69.

---

[1] This study was conducted by the second author under the direction of the first author in partial fulfillment of the requirements for the doctor degree in Clinical Psychology at Saint Louis University.

J. M. Oliver, Ph.D., is Professor of Psychology, Saint Louis University; Brett A. Hart, Ph.D., is a recent graduate of the doctoral training program in Clinical Psychology at Saint Louis University; Michael J. Ross, Ph.D., ABPP, is Professor of Psychology at Saint Louis University; and Barry M. Katz, Ph.D., is Professor of Research Methodology at Saint Louis University.

Attitudes Toward Perpetration of Sexual Abuse Against Children: A Study of Legal and Social Service Oriented Students

Scott A. Hubbartt and Sangeeta Singg
Angelo State University

Previous research has found attitudinal differences between legal and social service professionals regarding perpetration of sexual abuse against children. The question explored in this study is "Do these differences develop before or after a person enters a profession?" A sample of 192 college men and women, between 18 and 21 years of age, completed the Professional Attitudes Regarding Sexual Abuse of Children (PARSAC) Scale, which measured three factors: severity and seriousness of the issue, treatment versus punishment priority, and perpetrator identity. Based on a 2 (legal and social service interests) x 2 (men and women) factorial analysis of variance, it was concluded that certain biases exist before individuals enter their careers dealing with perpetrators of child sexual abuse. Women and men differed in their attitudes on all three factors, regardless of their professional orientation. The legal and social service oriented students differed mainly in their attitudes regarding the "treatment versus punishment" issue about the perpetrators.

While several studies (Parkinson, 1980; Saunders, 1988; Trute, Adkins, & MacDonald, 1992; Wilk & McCarthy, 1986) have revealed that attitudinal differences exist between legal and social service professionals toward those who perpetrate sexual abuse against children, no study to date has asked if these differences exist before individuals enter their professions. The major difference between the two types of professionals involves the "retribution versus rehabilitation" issue, legal service professionals viewing punishment as more effective and social service professionals opting for treatment ((Craft & Clarkson, 1985; Kelley, 1990; Saunders, 1988; Trute et al., 1992; Weekes, Pelletier, & Beaudette, 1995). This controversy also extends to sex differences (Attias & Goodwin, 1985; Deblinger, Lippmann, Stauffer, & Finkel, 1994; Eisenberg, Owens, & Dewey, 1987; Wagner, Aucoin, & Johnson, 1993; Wellman, 1993). In general, female professionals tend to be pro-treatment, while male professionals tend to be pro-punishment.

Author info: Correspondence should be sent to: Sangeeta Singg, Department of Psychology and Sociology, Angelo State University, P. O. Box 10907, San Angelo, Texas 76909 or via Internet (Sangeeta.Singg@angelo.edu).

North American Journal of Psychology, 2001, Vol. 3, No. 2, 243-252.

© NAJP

Trute et al. (1992) designed the Professional Attitudes Regarding the Sexual Abuse of Children (PARSAC) Scale and compared three of the key professional groups (police, child welfare, and community mental health workers) involved in the investigation and treatment of perpetrators of child sexual abuse. The PARSAC Scale measures three orthogonal factors.

(1) *Severity and seriousness of the issue* refers to the belief that child sexual abuse is widespread and psychologically affects the victims. (2) *Treatment versus punishment priority* taps into the philosophical position on punishment or treatment as the preferred intervention. (3) *Perpetrator identity* involves the belief that anyone can abuse a child and that perpetrators are not necessarily from deviant groups in society. Trute et al. (1992) found sex differences and differences between service sectors on all three factors. They found that men (compared to women) and police officers (compared to human service professionals) were less inclined to regard the sexual abuse as a widespread or serious problem, were more pro-punishment, and were more likely to see perpetrators as socially deviant individuals.

The question raised here is "Do the above attitudinal differences between two types of professionals develop before or after a person enters a career involving investigation/treatment of perpetrators of child sexual abuse?" In order to examine any pre-professional attitudinal differences, we compared young college men and women with future career interests in legal or social services. We also used the PARSAC Scale to measure three factors. Classification and age were controlled by using freshmen and sophomores between 18 and 21 years, because they were considered naive regarding the professional biases of their future careers.

In accordance with the study by Trute et al. (1992), we predicted that the social service orientation (SSO) group and women would have significantly higher mean scores (more lenient attitudes) than the legal service orientation (LSO) group and men on the three factors of the PARSAC Scale. Interactions between professional orientation and sex on the three factors were also of interest.

METHOD

Participants

A sample of 192 freshmen and sophomores (men = 96, women = 96; *M* age = 19.2, *SD* = .99) was obtained from a pool of 310 undergraduates enrolled in 15 classes at a regional southwestern university. On the basis of their future career interests (social service/legal service) and sex, they were divided into four groups: 49 SSO men, 54 LSO men, 47 SSO women, and 42 LSO women. The subject pool was 59% women, 71% Euro-American, 93% single, and 85% from middle class families (self-reported).

Instruments

The PARSAC Scale contains 14 items that are answered on a four-point Likert scale and measures professional attitudes regarding child sexual abuse (Trute et al., 1992). The three attitudinal domains measured by the PARSAC are as follows. (1) *Beliefs in regard to the extensiveness and seriousness of the issue* has six items (alpha = .71). A representative item from this subscale is "The media has blown sexual abuse out of proportion." This factor measures the tendency to see child sexual abuse as widespread and having important psychological impact on victims. (2) *Treatment versus punishment priority* has five items (alpha = .72). A sample item from this subscale is "The most effective intervention for child sex offenders is psychotherapy or counseling rather than jail." This factor measures the beliefs regarding the most effective professional intervention to deal with perpetrators. (3) *View regarding identity of those who perpetrate child sexual abuse* has three items (alpha = .61). A representative item from this subscale is "Women rarely sexually molest children." This factor measures the recognition that anyone can perpetrate sexual abuse against a child.

According to Trute et al. (1992), the PARSAC Scale addresses several erroneous public beliefs such as child sexual abuse rarely occurs, only socially deviant individuals molest children, and the belief that professionals are sensitive to the seriousness of child sexual abuse. This scale also taps into the current controversy concerning the most effective intervention to reduce sexually deviant acts of child molesters. A General Information Questionnaire (GIQ) containing demographic questions and questions about future career interest and college major was also used.

Procedure

The students from the 15 classes in the departments of psychology/sociology, history, and government were contacted and their voluntary participation was solicited. They were briefed about the survey and the careers under two orientations (e.g., legal service = police officers, lawyers, and probation officers; social service = counselors, child protective service workers, and social workers). Those who were majoring in the subjects they considered would lead to two types of careers were asked to complete the GIQ and PARSAC Scale. The goal was to obtain equal number of men and women who met the classification and age criteria.

RESULTS

Tables 1 and 2 present the results of the present study. As expected, the ANOVA results showed a significant main effect of professional orientation on Factor 2 (treatment versus punishment), with SSO group showing a significantly greater mean score (pro-treatment) than the LSO group. However, the main effects of professional orientation on Factor 1 (severity and seriousness) and Factor 3 (perpetrator identity) were not significant. The main effects of sex were significant on all three factors, women showing higher mean scores than men. No significant interaction between professional orientation and sex was found with regard to students' mean scores on any of the three factors.

TABLE 1 Mean Scores (Standard Deviations in italics) for Three PARSAC Scale Factors

	Orientation		
Factor	SSO	LSO	Total
Severity and Seriousness			
Men	19.90	19.89	19.90
	2.90	*2.20*	*2.56*
Women	20.82	20.98	20.90
	2.51	*2.08*	*2.32*
Total	20.36	20.43	
Treatment and Punishment			
Men	11.16	9.93	10.54
	2.66	*2.25*	*2.52*
Women	11.70	11.24	11.47
	2.20	*2.39*	*2.29*
Total	11.43	10.58	
	2.43	*2.40*	
Perpetrator Identity			
Men	14.65	14.57	14.61
	2.91	*1.85*	*2.45*
Women	15.35	15.17	15.26
	1.75	*2.20*	*1.96*
Total	15.00	14.87	
	2.39	*2.03*	

DISCUSSION

Because we wanted to compare the results of the present study of students preparing for their careers with a methodologically similar study of the professionals by Trute et al. (1992), we discussed our findings in reference to findings of their study. They compared professionals who were working with perpetrators of child sex abuse. Also, the use of PARSAC Scale and samples coming from large rural areas in both studies make the results more comparable.

Our findings differed from those of Trute et al. (1992) with regard to no significant differences between legal and social service oriented groups on factors 1 and 3 (severity and seriousness of the issue and perpetrator identity). However, we did find a significant difference between social and legal service orientation groups on Factor 2 (treatment versus punishment), which is in line with the results by Trute et al. (1992). Interestingly, the sex differences were found on all three factors in both studies.

TABLE 2 Analysis of Variance Results for Three PARSAC Scale Factors

Variable	*SS*	*F*
Severity and Seriousness		
Orientation (A)	.29	.05
Sex (B)	47.59	7.86**
A x B	.32	.05
Treatment and Punishment		
Orientation (A)	34.11	6.04**
Sex (B)	40.40	7.15**
A x B	6.90	1.22
Perpetrator Identity		
Orientation (A)	.82	.17
Sex (B)	19.84	4.02*
A x B	.14	.03

*$p < .05$; **$p < .01$

While Trute et al. (1992) found a significant difference between police officers and social service workers on Factor 1 (severity and seriousness). Their study included two groups of unequal size containing either mostly men or mostly women. When they controlled for sex on Factor 1, the professional groups lost their statistically significant difference. Therefore, it may be that both professionals and students are well aware of the widespread nature of child sexual abuse in our society and of the important psychological impact it can have on victims.

The present study and the study by Trute et al. (1992) coincided in their findings on Factor 2 (treatment versus punishment). One explanation why a significant difference existed between professional orientation groups of students on Factor 2 in the present study stems from Holland's (1997) vocational theory. Holland demonstrated that occupational personality types exist and persons with certain social biases are attracted to certain career paths. They are products of their environments, which influence their personal orientations and eventual career choices. Thus, persons bring certain biases to their professions.

Results for Factor 3 (perpetrator identity) in the present study imply that both the legal and social service orientation groups held similar views about the family backgrounds and identity of child molesters. In other words, they recognized that anyone could perpetrate sexual abuse against a child and offenders are not necessarily from a deviant group. This was not the case in the study of professionals by Trute et al. (1992). Their study showed that police officers identified perpetrators as coming from deviant sectors of society. A logical explanation for our finding could be that the participants in the present study were not closely involved with victims or perpetrators of child sexual abuse as were the police and social service workers in the Trute et al. (1992) study.

Similar to the study by Trute et al. (1992), women in the present study scored significantly higher than men on all three factors. This is consistent with the findings of another recent study of students by Wellman (1993). The results for factor 1 (severity and seriousness) indicate that college women consider child sexual abuse to be a more serious and prevalent issue than do college men. This perception of college women concerning the serious nature of child sexual abuse may exist because when sexual abuse occurs, the victims are mostly female (Finkelhor, Hotaling, Lewis, & Smith, 1990). Hence, it is reasonable to suspect that during their youth, females may be exposed to sexual abuse experiences more than males.

Another reason could be that mens' views toward sex might be diminished by society's seeming approval of teenage boys who engage in sexual intercourse with older women. When Broussard, Wagner, and Kazelskis (1991) studied the impact of perpetrator's sex and victim's sex, the undergraduate college students viewed sexual intercourse of a female perpetrator (age 35) with a 15-year old male victim as less indicative of child sexual abuse.

Briere and Runtz (1989) presented another possible explanation for why men have less serious attitudes toward child sexual abuse. They collected data from a sample of 193 male undergraduate college students. Their findings suggested that sexual response to children may be more common in our society among men than we realize. Twenty-one percent admitted being sexually attracted to some small children (the age range of small children was not specified). Nine-percent reported fantasizing about having sex with a child and 5% admitted to masturbating at least once to such fantasies. Seven-percent of the students indicated some likelihood of engaging in sex with a child if they could avoid detection or punishment. Findings of this nature are very alarming, and such studies need to be replicated.

In line with the study by Trute et al. (1992), the results of the present study showed that college women tend to have less punitive attitudes toward sex offenders than men (Factor 2, treatment versus punishment). This attitude may reflect the female nurturing and caretaking role in our society.

Again, in keeping with Trute et al. (1992), men seemed to accept the myth that child sexual abuse occurs infrequently and is initiated by socially abnormal individuals (Factor 3, perpetrator identity). This belief may exist as a form of denial on the part of men, because the majority of sex offenders are male (Finkelhor et al., 1990). Furthermore, Finkelhor et al. and other researchers (Bolen, 2000; Cupoli & Sewell, 1988) have shown that a male perpetrator is usually a member of the family or someone known to the child and is rarely a stranger (the traditional stereotype). If a man was to accept that perpetrators are from the normal population, this may in part be a personal attack on his dignity, indicating that he is also capable of being connected to this ominous subgroup.

Some limitations of the present study are worth noting. Some possible intervening variables were not controlled, such as exposure to information about child sexual abuse, possible child sexual abuse history of the participants, and ethnicity. Also, career interests of students could be subject to change as they progress in college. There was no measure of how strongly a student was interested in the social or legal service type of profession.

The findings of the present study suggest that young men and women differ in their attitudes toward perpetration of sexual abuse against children, regardless of their professional orientation. Thus, the sex difference might be a function of various belief systems that also influence people into choosing different career paths. This difference develops before individuals enter their careers. Therefore, it may be fair to include a balance of male and female jurors when there is a decision to be made as to what should be done with a perpetrator of child sexual abuse.

The only difference between the professional orientation groups in the present study has to do with the issue of "retribution versus rehabilitation." The implication is that this difference does not develop after one joins either of the two professional groups described here; rather it may pre-exist, indicating certain personal characteristics of people that group them into different pre-occupational types. For more objectivity and later cohesive effort of communication among professionals, maybe these differences should be addressed during the training programs in college. What to do with the perpetrators of child sexual abuse is just as serious an issue as the abuse itself. It deserves attention from all relevant professional sectors so that the most effective measures can be taken to treat the source of the problem, the perpetrator. If the leaders and professionals of our society do not seriously focus on this aspect of the problem, the cause of the problem will never be targeted. Abel and Rouleau (1995, p. 150) clearly expressed this point of view when they concluded that "the path to no victims is to treat the perpetrators."

REFERENCES

Abel, G. G. & Rouleau, J. L. (1995). Sexual abuses. *The Psychiatric Clinics of North America, 18* (1), 139-153.

Attias, R., & Goodwin, J. (1985). Knowledge and management strategies in incest cases: A survey of physicians, psychologists, and family counselors. *Child Abuse and Neglect, 9*, 527-533.

Bolen, R. M. (2000). Extra-familial child sexual abuse: A study of perpetrator characteristics and implications for prevention. *Child Abuse and Neglect, 6*, 1137-1169.

Briere, J., & Runtz, M. (1989). University males' sexual interest in children: Predicting potential indices of pedophilia in a non-forensic sample. *Child Abuse and Neglect, 13*, 65-75.

Broussard, S., Wagner, W. G., & Kazelskis, R. (1991). Undergraduate students' perceptions of child sexual abuse: The impact of victim sex, perpetrator sex, respondent sex, and victim response. *Journal of Family Violence, 6* (3), 267-278.

Craft, J., & Clarkson C. (1985). Case disposition recommendations of attorneys and social workers in child abuse investigations. *Child Abuse and Neglect, 9*, 165-174.

Cupoli, J. M. , & Sewell, P. M. (1988). One thousand fifty-nine children with a chief complaint of sexual abuse. *Child Abuse and Neglect, 12*, 151-162.

Deblinger, E., Lippmann, J., Stauffer, L., & Finkel, M. (1994). Personal versus professional responses to child sexual abuse allegations. *Child Abuse and Neglect, 18* (8), 679-682.

Eisenberg, N., Owens, R., & Dewey, M. (1987). Attitudes of health professionals to sexual abuse and incest. *Child Abuse and Neglect, 11* (1), 109-116.

Finkelhor, D., Hotaling, G., Lewis, L. A., & Smith, C. (1990). Sexual abuse in a national survey of adult men and women: Prevalence, characteristics, and risk factors, *Child Abuse and Neglect, 14*, 19-28.

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environment* (3rd ed.). Odessa, FL: Psychological Assessment Resources, Inc.

Kelley, S. J. (1990). Responsibilities and management strategies in child sexual abuse: A comparison of child protective workers, nurses and police officers. *Child Welfare, 69*, 43-51.

Parkinson, G. (1980). Cooperation between police and social workers: Hidden issues. *Social Work, 25*, 12-18.

Saunders, E. (1988). A comparative study of attitudes toward child sexual abuse among social work and judicial system professionals. *Child Abuse and Neglect, 12*, 83-90.

Trute, B., Adkins, E., & MacDonald, G. (1992). Professional attitudes regarding the sexual abuse of children: Comparing police, child welfare and community mental health. *Child Abuse and Neglect, 16*, 359-368.

Wagner, W., Aucoin, R., & Johnson, J. (1993). Psychologists' attitudes concerning child sexual abuse: The impact of sex of perpetrator, sex of victim, age of victim, and victim response. *Journal of Child Sexual Abuse, 2* (2), 61-74.

Weekes, J., Pelletier, G., & Beaudette, D. (1995). Correctional officers: How do they perceive sex offenders? *International Journal of Offender Therapy and Comparative Criminology, 39* (1), 55-61.

Wellman, M. M. (1993). Child sexual abuse and gender differences: Attitudes and prevalence. *Child Abuse and Neglect, 17* (4), 539-547.

Wilk, R. J., & McCarthy, C. R. (1986). Intervention in child sexual abuse: A survey of attitudes. *Social Casework: The Journal of Contemporary Social Work, 67* (1), 20-26.

Responsibility for this article is shared equally by the authors, whose names are listed alphabetically.

Self-protective Pessimism: Optimistic Bias in Reverse

John Chapin
Penn State University - Beaver

Over 100 published studies in the past 20 years confirm the presence of unrealistic optimism in a variety of contexts, particularly health risk perception. This study used optimistic bias measures to consider the effect of pessimism on health risk perception and risk-taking behaviors. The findings of a survey of a suburban community sample indicate that study participants exhibited pessimistic bias regarding the likelihood of reaching their life goals, and that this pessimism was linked to age/experience. Younger participants tended to exhibit a pessimistic view of the future, which in turn appears to encourage certain risk-taking behaviors, including alcohol abuse, vehicular speeding, unprotected sexual activity, and cigarette smoking.

Weinstein (1980) coined the phrase "optimistic bias" to describe the common misperception that "bad things happen to other people, but not to me." Weinstein argued that individuals make comparative risk assessments in an egocentric manner, paying little attention to the risk status of others when asked to determine their own relative risk. In lay terms, individuals believe they are less vulnerable to risks than others.

Optimistic bias is a robust finding and has been replicated in a variety of contexts, including HIV/AIDS risk (Ellen, Boyer, Tschann & Shafer, 1996; Harris, 1996), sexually transmitted disease (STD) risk (Kaplan & Shayne, 1993; Turner, 1993), pregnancy risk (Eldridge, Lawrence, Little, Shelby & Brasfield, 1995; Smith, Gerrard, & Gibbons, 1997), cancer risk (Aiken, Febaughty, West, Johnson, & Luckett, 1995; Fontaine & Smith, 1995), smoking risk (Strecher, Kreuter & Korbin, 1995), substance abuse risk (Hansen, Raynor, & Wolkenstein, 1991; Miller, 1991), and general health risks (Glanz & Yang, 1996; Hoorens, 1996).

While the concept of optimistic bias has generated a great deal of research, only one study (Sultan & Bureau, 1999) considered another possibility… pessimistic bias. What effect does a pessimistic viewpoint have on risk-perception and attitudes? Do those who believe they are especially prone to negative outcomes take more self-protective precautions or passively accept their fate? The purpose of the current study is to apply knowledge gained through the optimistic bias literature

---

Author info: Correspondence should be sent to: Dr. John Chapin, Penn State U., 110 Michael Baker Bldg., Monaca, PA 15061.

North American Journal of Psychology, 2001, Vol. 3, No. 2, 253-262.
© NAJP

to a pessimistic sample, thereby gaining insight on health risk-perception and risk-taking behaviors.

Over 100 studies published since 1980 confirm the presence of optimistic bias and explore the impact of such misperceptions on individuals (see Weinstein 1987, 1989 for reviews). The over-reliance on college student samples in optimistic bias research has been critiqued frequently (Weinstein, 1987, p. 483):

> College students are a biased sample of the population in other respects. They are likely to be healthier, to become better educated, and to come from wealthier homes than the average person their age. Thus, they may be correct when they claim that their chances of experiencing harm are less than average (p. 483).

Weinstein (1987) sought to control for this bias by conducting a phone survey of residents of a community in New Jersey. He still found optimistic bias regarding a variety of health risks, including drug addiction, attempted suicide, senility, diabetes, and gum disease. Unfortunately, this "community sample" excluded anyone under the age of 18. Chapin's (2000) study of New Jersey youth found adolescents to be equally (if not more) optimistic by comparison with previous samples, especially regarding sexual risks. The critique of the over-reliance on college students is valid. If adolescents tend to be more optimistic than college students regarding health risks, then age may be positively correlated with pessimism.

Weinstein (1982,1987) concedes that experience has great influence on optimistic bias. Individuals use past experiences to predict their future vulnerability, mistakenly believing that if it hasn't happened yet, it won't happen:

> Such a belief may arise when people believe the problem has a hereditary basis and will appear early in life if it is going to appear at all (e.g. juvenile diabetes and asthma). In other cases (e.g., tooth decay) people believe that vulnerability is a constitutional matter, so if the problem has not appeared, their bodies must be resistant. Furthermore, some problems may seem to be caused by one's behavior or personality (e.g., obesity and drug addiction), and people may conclude that the absence of a problem at their age means they do not have the weakness of character that allows it to develop (Weinstein, 1987, p. 484).

Numerous studies confirm a link between experience and optimistic bias (McCoy, Gibbons, Gerrard, & Suftka, 1992; Moore & Rosenthal, 1991). Age alone, then, may fail to predict optimistic bias, depending on the outcome of experiences. If an individual speeds frequently, but fails to get a ticket, that individual is likely to be optimistic that he/she has a special ability to avoid police, and will continue to speed. On the other hand, if a person speeds frequently and receives numerous tickets, would he or she become pessimistic, now believing that he/she is some sort of "police magnet?"

While the influence of unrealistic optimism on health outcomes and risk-behaviors has been broadly investigated, less is known about those prone to pessimism. Norem (1986) argued that anxiety is used as a personal motivator. Results of Norem's experiment indicated that pessimism (in the form of low expectations) coupled with high anxiety negatively impact individual perceptions when faced with risky situations. Likewise, Donegal (1999) found pessimism to be predictive of performance deterioration and accuracy in experimental tasks. Others confirm a link between pessimism and failure (Mikolajczyk, 1999) and impaired judgment (Keltner, 1993).

The first study to use Weinstein's (1980) optimistic bias instrument to examine pessimism found "unrealistic optimism" to be a correlate of dispositional optimism (Sultan & Bureau, 1999). A respondent's place on a continuum between optimism and pessimism was found to predict vulnerability appraisals regarding sexual risk-taking situations. Highly optimistic students believed they were less vulnerable to sexual risks, while highly pessimistic students believed they were more prone to sexual risks. Sultan and Bureau sought to link the pessimism literature to the optimistic bias literature; however, their findings suggest a little pessimism may be a healthy thing. The current study sought to extend these findings by avoiding the biases inherent in the overuse of college student samples and through the use of multiple risk-behaviors.

Based on the literature I hypothesized that the persons in the present sample would be less likely than other people their age in the U.S. to believe they would reach their life goals (pessimistic bias). Furthermore, pessimistic bias will increase as age increases, and as pessimistic bias increases, the number of risks taken will decrease. Finally, as pessimistic bias increases, the frequency of risks will decrease.

METHOD

Participants

Participants were randomly selected from several communities in Allegheny County, Pennsylvania. Allegheny County suffers one of the highest poverty rates in Pennsylvania (21%), with an average per capita income of $15,115 (Bureau of the Census, 1994). Pittsburgh is one of only four American cities that experienced population loss in the 1990s. The largely Euro-American (87%) population (African-Americans = 11%) reside in urban (96%) or suburban environments in or near Pittsburgh (Pennsylvania State Data Center, 1999). The Pittsburgh area provides an ideal location for this study. One hundred sixty two residents anonymously completed the survey instrument for a nominal payment. The sample was somewhat skewed, over-representing females (61%) and Euro-Americans (99%) ranging in age from 14 to 78 (*M* = 34.9, *SD* = 16.1).

Procedure

Households were randomly selected from local telephone directories. People who received calls were informed about the survey and asked if they and/or other members of the household (ages 14 and up) would participate. Calls continued until 200 individuals agreed to have surveys mailed to their home (23% success rate). Of the individuals who agreed to participate by phone, 81% (162) completed consent forms and surveys and returned them by mail.

Measures

Optimistic bias regarding health risks was measured with a standard instrument designed by Weinstein (1980). The procedure asked participants to compare their relative risk of four target variables (vehicle accident, AIDS, STD, cancer), with "other people my age in the U.S." Comparative risk assessment was measured on a 7-point scale (-3 = "much less" than other students in the USA, +3 = "much greater" than other students in the USA). Optimistic bias was indicated by a group mean significantly less than zero, and pessimistic bias by a group mean significantly greater than zero. A mean of zero indicated no bias. On Table One these four were combined to yield one score called "pessimism: health risks."

Pessimistic bias regarding life goals was measured with a similar instrument, which replaced the target risk with a positive statement: "Compared to other people my age in the U.S., my chances of reaching my life goals are…" The same response scale was used; however, the scale was reversed for consistency across the optimism and pessimism measures. In Table One this is called "pessimism: life goals."

Consistent with previous reviews (Weinstein, 1987), the optimistic bias instrument is highly reliable (α = .88); however, two threats to validity have been suggested. Because young people are more likely than older people to believe the future will be better than the present, the overuse of college student samples is problematic. The second threat to validity is the failure to measure "actual risk." The current study addressed the first concern by utilizing a community-based sample. All attempts at measuring one individual's "actual risk" of a particular risk (i.e. contracting a STD from unprotected sex) ultimately fail, because they must rely on percentages and estimates. Couching discussion of optimism/pessimism at the group level minimizes, but doesn't eliminate, the threat to validity.

In addition to demographic information (age, gender, geographic location), individuals self-reported the age of initiation of sexual activity. Several domains of risk behavior were measured using an instrument adapted from Arnett (1996). Participants indicated past risk by reporting whether they had *ever done* any of the following: smoked cigarettes, had unprotected sex (without a condom), had protected sex (with a condom), driven 10-miles over the speed limit. Present risk was measured by asking participants to estimate the number of days they had engaged in each of the four risk-behaviors in the past 30 days. Five response categories were provided: none, 1-2 days, 3-10 days, 11-20 days, almost every day.

RESULTS

Hypothesis 1 predicted pessimistic bias among the respondents. The positive mean (*M* = .69, *SD* = 1.4) indicated that study participants exhibited the tendency to believe they were less likely than other people their age in the U.S. to achieve their life goals. A single-sample-t-test was used to test the hypothesis that the mean of pessimism was significantly greater than zero. As predicted in H1, participants exhibited pessimistic bias, *t* (161) = 6.2, *p*< .001.

Table 1 Zero-Order Correlations Among Pessimistic Bias and Study Variables

	2	3	4	5	6	7	8
1. Pessimism (LG)	- 32**	-31**	-25**	-12	-06	-05	-02
2. Age	--	54**	07	01	-03	10	-09
3. # of Risks		--	12	29**	35**	43**	26**
4. Pessimism (HR)			--	05	06	-04	03
5. Freq: Drinking				--	22*	28**	21*
6. Freq: Speeding					--	18	08
7. Freq: Unprot. Sex						--	22*
8. Freq: Smoking							--

Note. Because optimistic bias has a negative mean, all signs in row 4 should be reversed for interpretation. All decimals omitted. LG = life goals. HR = health risks.
*p< .01. **p<.001.

Table 1 summarizes findings related to Hypotheses 2 through 4. As predicted in H2, pessimistic bias was related to age, at least for life goals; however, the relationship was in the opposite direction. Pessimism appears to decrease with age. Weinstein (1989) argues that optimistic bias decreases with direct experience with a target variable; i.e. if one drinks and drives and experiences a negative consequence (an accident), one becomes less optimistic about avoiding accidents while under the influence. Perhaps in this case, participants in Allegheny County began with low expectations regarding the future and became more optimistic with age as they experienced positive outcomes (success).

Hypotheses 3 and 4 explored the effect of pessimism on current and past risk behaviors. As predicted in H3, an inverse relationship appeared between life goals pessimistic bias and the number of health risks taken (alcohol consumption, speeding, unprotected sex, and cigarette smoking). People who are pessimistic about achieving their life goals take fewer types of risks than optimistic people.

TABLE 2 Summary of Linear Regression Analysis for Variables Predicting Pessimistic Bias

	Pessimistic bias Adj $r^2 = .11$ n = 153		
Predictor	B	SE B	β
Age	- .02	.01	-.22**
# of risks	- .21	.10	-.18*

*p< .05 **p< .01.

Standard multiple regression was used to identify the best predictors of pessimistic bias. The results are displayed in Table 2. Analysis of residual plots indicates that assumptions regarding normality, linearity, and homoscedasticity were met. Both age and experience (number of risks experienced) remained in the model. Age was the best predictor of pessimistic bias, with pessimistic bias decreasing with age. While the results suggest changes in currently held assumptions about optimistic/pessimistic bias, the results of this analysis should be interpreted with caution due to the low adjusted *r*-square.

DISCUSSION

Contrary to H4, no relationship was found between pessimistic bias and the frequency of the four risk behaviors during the past 30-day period. It is interesting to note, however, that frequency of risks was significantly related to the number of life goal risks. This suggests the need for further exploration of the relationship between pessimistic bias and risk behaviors, and also lends support to Jessor and Jessor's (1977) argument that "problem behaviors" should be studied in groups, because individuals tend to engage in multiple risk behaviors simultaneously. The intercorrelations on Table 1 among the four risk behaviors also support this conclusion.

Optimistic bias is a robust finding. Over 100 published studies in 20 years confirm the presence of unrealistic optimism in a variety of contexts, particularly health risk perception. This study is only the second to use optimistic bias measures to consider the effect of pessimism on health risk perception and risk-taking behaviors. The study also contributes to the literature by avoiding the over-reliance on college student samples by focusing on an urban community ranging in age from 14 to 78.

The findings indicate that study participants exhibited pessimistic bias regarding the likelihood of reaching their life goals, and that this pessimism was linked to age/experience. Younger participants tended to exhibit a fatalistic view of the future, which in turn appears to discourage certain risk-taking behaviors, including alcohol abuse, vehicular speeding, unprotected sexual activity, and cigarette smoking. These behaviors in particular have been associated with adolescents who exhibit high degrees of optimistic bias (Chapin, 2000). This finding seems to suggest that a little self-protective pessimism may promote healthy choices and behaviors. As individuals experience success, the degree of pessimistic bias tends to decline, yielding a more optimistic outlook (in some cases unrealistically so), and the adoption of more risky habits.

Such findings are hopeful in that they suggest a strategy for health campaigns that emphasize personal risk over high-risk groups and national statistics. People who believe they are prone to negative outcomes can and do take precautions. Building on the existing literature, the results of this study indicate that people who take health risks tend to take multiple risks and take them more frequently.

The optimism/pessimism continuum presents a promising framework for understanding how people process health risk information, which people are prone to risk-taking behaviors, and how to encourage people to adopt self-protective behaviors. The use of a single measure encompassing the continuum would clarify relationships suggested in the current study, and may be more applicable to specific risk behaviors. Future research could also emphasize the formation of optimistic/pessimistic bias and explore methods of priming either response through health promotion campaigns.

While the results of the current study are based on a random sample, it intentionally over-represents urbanites and suburbanites in an economically depressed geographic area. It also over-represents Euro-Americans. Most optimistic bias studies, which included items about life goals, found people to be unrealistically optimistic about the future. This may be due to the over-reliance on college student samples of previous studies; however, one study does not provide a basis for drawing conclusions about the source of optimism or pessimism regarding the future.

REFERENCES

Aiken, L., Fenaughty, A., West, S., Johnson, J. & Luckett, T. (1995). Perceived determinants of risk for breast cancer and the relations among objective risk perceived risk, and screening behavior over time, *Women's Health, 1,* 27-50.

Bureau of Census (1994). Pennsylvania. Harrisburg.

Chapin, J. (2000). Third-person perception and optimistic bias among urban, minority, "at-risk" youth. *Communication Research, 27* (1), 51-81.

Donegal, L. (1999). Pessimism as a predictor of performance deterioration after false negative feedback. *Dissertation Abstracts International, 60* (2B), 0824.

Eldridge, G., Lawrence, J., Little, C., Shelby, M & Brasfield, T. (1995). Barriers to condom use and barrier method preferences among low-income African-American women, *Women and Health, 23* (1), 73-89.

Ellen, J., Boyer, C., Tschann, J & Shafer, M. (1996). Adolescents' perceived risk for STD and HIV infection, *Journal of Adolescent Health, 18* (3), 177-181.

Fontaine, K. (1994). Effect of dispositional optimism on comparative risk perceptions for developing AIDS, *Psychological Reports, 74,* 843-846.

Glanz, K & Yang, H. (1996). Communicating about the risk of infectious diseases, *Journal of the American Medical Association, 275* (3), 253.

Hansen, W., Raynor, A. & Wolkenstein, B. (1991). Perceived personal immunity to the consequences of drinking alcohol: The relationship between behavior and perception, *Journal of Behavioral Medicine, 14,* 205-225.

Harris, P. (1996). Sufficient grounds for optimism: The relationship between perceived controllability and optimistic bias, *Journal of Social and Clinical Psychology, 15*(1), 9-52.

Hoorens, V. (1996). Sufficient grounds for optimism? The relationship between perceived controllability and optimistic bias, *Journal of Social and Clinical Psychology, 15*(1), 9-52.

Jessor, R. & Jessor, S. (1977). *Problem behavior and psychosocial development: A longitudinal study of youth.* New York: Cambridge University Press.

Kaplan, B. & Shayne, V. (1993). Unsafe sex: Decision making bias and heuristics. *US AIDS Education and Prevention*, 5(4), 294-301.

Keltner, D. (1993). Beyond simple pessimism: Effects of sadness and anger on social perception. *Journal of Personality and Social Psychology, 64* (5), 740-752.

McCoy, S., Gibbons, F., Reis, T., Gerrard, M., Luus, E., & Sufka, A. (1992). Perceptions of smoking risk as a function of smoking status, *Journal of Behavioral Medicine, 15* (5), 469-488.

Mikolajczyk, M. (1999). Optimism-Pessimism and students' plans and professional expectations. *Psychologia Wychowawcza, 42* (1), 11-26.

Miller, C. (1991). Risk perception: An investigation of its accuracy and impact, *Dissertation Abstracts International*, 52 (01B), 523.

Moore, S. & Rosenthal, D. (1991). Adolescent invulnerability and perceptions of AIDS risk, *Journal of Adolescent Research, 6*(2), 164-180.

Norem, J. (1986). Defensive pessimism: Harnessing anxiety as motivation. *Journal of Personality and Social Psychology, 51* (6), 1208-1217.

PA State Data Center (1999). Harrisburg. PA. (author)

Smith, G., Gerrard, M. & Gibbons, F. (1997). Self-esteem and the relation between risk behavior and perceptions of vulnerability to unplanned pregnancy in college women, *Health Psychology, 16*(2), 137-146.

Strecher, V. Kreuter, M. & Kobrin, S. (1995). Do cigarette smokers have unrealistic perceptions of their heart attack, cancer an stroke risks, *Journal of Behavioral Medicine, 18*(1), 45-54.

Sultan, S. & Bureau, B. (1999). Which optimism in health psychology? *European Review of Applied Psychology, 49* (1), 43-51.

Turner, J. (1993). Teenage sexual risk-taking and perceived invulnerability: The influence of adolescent egocentrism, *Dissertation Abstracts International, 5* (07A), 2752.

Weinstein, N. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology, 39*, 806-460.

Weinstein, N. (1982). Unrealistic optimism about susceptibility to health problems. *Journal of Behavioral Medicine, 5*, 441-460.

Weinstein, N. (1984). Why it won't happen to me: Perceptions of risk factors and susceptibility, *Health Psychology, 3*(5), 431-457.

Weinstein, N. (1987). Unrealistic optimism about susceptibility to health problems: Conclusions from a community-wide sample. *Journal of Behavioral Medicine, 10*(5), 441-458.

Weinstein, N. (1989). Optimistic bias about personal risks, *Science, 246* (4935), 1232.

The author gratefully acknowledges the contribution of Troy Hough (Penn State University) to this paper.

Lost Letter Technique: Returned Responses to Battered and Abused Women, Men, and Lesbians

F. STEPHEN BRIDGES
C. BENNETT WILLIAMSON
PAUL C. THOMPSON
MARGARET A. WINDSOR

The University of West Florida

Of 840 letters "lost" in the Florida Panhandle, 467 (55.6%) were returned in the mail by the finders (the altruistic response). Addressees' affiliations were associated with significantly different rates of return; fewer emotive Advocates for Battered and Abused Lesbians (ABAL) addressees were returned than controls. Hierarchical loglinear models were utilized to uncover the complex relationships between geographic location (Ohio and Florid/Alabama), community size (city and town), addressee affiliation (controls and emotive homosexual affiliate) and returned response (returned and not returned). The results were significant and suggest that letters were more likely to be returned if they were: lost in Ohio than in Florida/Alabama; lost in a city rather than in a small town; addressed to control affiliates than emotive lesbian ones; or were addressed to a person than an organization. These findings appear to support the basic premise of Milgram's lost letter technique, namely, that the probability of lost letters being returned depends on the social and political attributes of the addressees' affiliations.

Persons asked to complete surveys or participate in interviews often recognize that their responses will be studied. Participants may have concern about how they are evaluated which can in turn have a strong effect on what they say, especially regarding socially and politically sensitive subjects. Researchers recognized this problem and developed another way of measuring community attitudes, the lost letter technique. They reported this technique to be useful as a research tool for measuring public sentiment toward political organizations, including deviant ones (Milgram, Mann, and Harter, 1965). These researchers addressed stamped letters to fictitious persons and groups and dispersed a large number of them in public places. It was assumed that a passerby would either ignore the letter, respond to it, but not take it, or pick it up and take

---

Author info: Correspondence should be sent to: F. Stephen Bridges, Ed.D., Division of Health, Leisure and Exercise Science, U. of West Florida, 11000 University Parkway, Pensacola, FL 32514-5750, e-mail (fbridges@uwf.edu).
North American Journal of Psychology, 2001, Vol. 3, No. 2, 263-276.
© NAJP

it with the intention of returning it in the mail. Presumably, the attitude of the finder toward the particular addressee (person or group) might affect this decision. Thus, the dependent variable was the number of "lost" letters returned in the mail. Milgram (1970) hypothesized that in cities people help others less often than in small towns. Other researchers tested this hypothesis and addressed letters to fictitious persons who differed in evident conformity to conventional social norms (Hansson & Slade, 1977; Whitehead & Metzger, 1981). Neither study found a difference in the urban rate of return of lost letters regardless of whether the addressee was deviant, whereas from small towns fewer lost letters were returned if the addressee was deviant.

The willingness of Americans to extend civil liberties to any number of unpopular groups has interested social scientists for many years. Hunter (1987) wrote about the ever-increasing cultural polarization of American society since the mid-1970s. For about fifty years, the research has shown Southerners to be more reluctant than non-southerners to extend civil liberties to various unpopular groups (Abrahamson & Carter, 1986; Ellison & Musick, 1993; Middleton, 1976; Nunn, Crockett, & Williams, 1978; Stouffer, 1955; Tuch, 1987; Wilson, 1986). Low levels of tolerance (i.e., the lack of willingness to extend civil liberties to deviant groups) continue to persist in the south due in part to protestant fundamentalism (Borg, 1997; Ellison & Musick, 1993). In Nunn, Crockett, and Williams's (1978) study (as cited in Ellison & Musick, 1993) the south, as a regional subculture, was unfavorably characterized as follows: "Localism, defined as a marked preference for and identity with the locale of one's birth, is joined in the South with an insularity of mind that is slow to change, actively belligerent toward the new, and openly intolerant toward a diversity of viewpoints." (p. 380). Surveys of public opinion in the 1980s showed widespread fear of AIDS, with much inaccurate information about how the disease is spread, and a public willingness to support Draconian policies that would restrict civil liberties for the purpose of battling the disease (Herek, 1999). Herek (1999) defined AIDS-related stigma as "prejudice, discounting, discrediting, and discrimination directed at people perceived to have AIDS or HIV, and the individuals, *groups*, and communities *with which they are associated*" (p. 1103). Historically in America, symbolic AIDS stigma has focused primarily on male homosexuality, and many Americans continue to equate AIDS with this sexual orientation (Herek, 1999; Herek & Capitanio, 1999).

"Person-positivity" bias is the general tendency to evaluate any individual social objects, not just self, more positively than groups or collections of individuals (aggregates) (Sears, 1983). It is based on Sears' (1983) three assumptions about the uniqueness of persons as objects evoking attitudes (e.g., animate objects such as people ranging to inanimate ones, like rocks or a desert) and why they attract relatively more positive response from other persons. First, all objects evoking attitudes could be conceptualized as representing different points on a "personhood" dimension, with specific individuals at one end, collections of individuals (aggregates) in the middle, and group stereotypes somewhere between aggregates and inanimate objects at the other impersonal end. In fact, even animals and other beings can vary a lot on this "personhood" dimension, i.e., machines and animal characters in children's stories were evaluated more positively when given human names than when identified in the usual nonhuman manner. Thus, the characterization of the dimension as "personhood" stresses the concept of similarity between any given object and a human person (Sears, 1983). Secondly, Sears (1983) assumed that the more an object resembles a person, the more evaluators will perceive its similarity to themselves, and the more favorably it would be evaluated. But he noted that even groups or aggregates can vary a lot on the "personhood" dimension: for example, group objects that bring to mind ordinary persons, to whom an evaluator might feel more similar, would be evaluated more positively than group objects that characterize an abstract type. Kaplan (1976) had noted the universal existence of a preference for a "positivity" bias for "people in general," but he also indicated that marked individual differences exist. Sears' (1983) third assumption is that generalized attitudes toward different objects on the "personhood" dimension tend to be compartmentalized, even when they seemed to be closely related to each other. In sum, the general tendency to evaluate persons more positively than groups is because similarity enhances liking (Berscheid & Walster, 1978; Byrne, 1971). Sears (1983) and his colleagues (Sears, Brown, & Ditto, 1982; Sears & Whitney, 1973) used this well-documented relationship between similarity and liking to explain why evaluations of social groups (e.g., politicians in general) are often less positive than evaluations of specific individual members of those groups (e.g., specific US Senators).

Some researchers (Milgram, *et al.*, 1965; Simmons & Zumpf, 1983; Wicker, 1969) have reported a higher rate of return for personal letters than for those letters addressed to political groups and other institutions. In addition, higher rates of return were reported for lost letters addressed to: 1) a person rather than letters addressed to a person(s) "in care of" a group(s) (Bridges & Coady, 1996; Bridges, Ryan, & Scheibe, 1998; Bridges, Williamson, & Scheibe, 1998; Hansson & Slade, 1977; Whitehead & Metzger, 1981), 2) a person than to a group's name followed underneath by "Attention to: person's name" (Weiner, 1975), 3) a person than those letters addressed to a person "in care of" a political group. These in turn had a higher rate of return than those letters addressed to political groups (Jacoby & Aranoff, 1971). Simmons and Zumpf (1983) offered Sears' hypotheses as a reasonable explanation for their higher return rate across conditions (kinds of addressees) for lost letters addressed to individual persons rather than to groups (e.g., committees) and for their high overall rates of return of letters.

Sears (1983) explained that the reason why machines and animal characters with human names were evaluated more positively was because the application of names changed their positions on the "personhood" dimension. We applied this concept by adding a person's name to addresses for lost letter return, so as to influence the position of the groups' stereotypes on the "personhood" dimension, thereby allowing them to be evaluated more favorably. Drawing from the work of Sears (1983) and Jacoby and Aranoff (1971) the present experiment sought to use a person-positivity effect, i.e., more letters are returned when they are addressed to a 'person in care of a group' than to a just a 'group', to improve upon the lost letter technique as a method for assessing community responses about a somewhat political issue. We predicted, after considering Kaplan's work and his modification of Sears' similarity hypothesis, that any "person-positivity" effect should produce both a *higher return rate* across kinds of addressees for lost letters addressed to a 'person in care of a group' than to a 'group' alone and a *higher overall rate of return of letters*.

The purposes of this study were to: (1) determine if there is a significant relationship among return response (letter returned/not returned), geographic location, community size, and addressee affiliation for residents of cities and small towns in Ohio, Florida, and Alabama, (2) to determine to what extent the variables (addressee affiliation or geographic location or community size) influence return rate, and (3) to determine if "person-positivity" would influence return rates of lost letters.

METHOD

Lost Letters

The present study modified previous designs of others (Bridges, 1996; Bridges, Anzalone, & Anzalone, unpublished) by adding new affiliations for addressees and by adding other conditions such as community size, person-positivity effect, and geographic location. The affiliates for two of the addressees were designed to represent groups that in the South do not arouse very high levels of emotional involvement. In contrast, the affiliate for the third addressee was designed to represent an unpopular and often stigmatized homosexual group. Thus, the affiliations were renamed either Advocates for Battered and Abused Women

(ABAW), Advocates for Battered and Abused Men (ABAM), and Advocates for Battered and Abused Lesbians (ABAL). The geographic location allowed comparison of the lost letter technique at two culturally and geographically different test sites. The city of Pensacola in Florida and surrounding small towns in adjacent Alabama are culturally dissimilar to the city of Lima and the nearby small towns of Ada and Kenton which are located far away in the northwestern corner of Ohio. The cultural differences between these geographic locations, i.e., Southern versus Northern, might influence response to addressees' affiliations and return of lost letters. Thus, the geographic location condition referred to whether a lost letter was dropped in the South in the city of Pensacola, Florida and surrounding small towns in adjacent Alabama or in the North in the city of Lima, Ohio and in nearby small towns of Ada and Kenton. The "person-positivity" condition referred to whether a lost letter was addressed to a "person in care of a group" or just a "group."

Procedure

Two field experiments used 840 "lost letters." In the first experiment, a total of 540 lost letters, i.e., either 50 letters (only 40 for small towns) for each of three affiliations for addressees in two sizes of community and for two conditions that manipulated "person-positivity," were distributed in public places in Florida and Alabama: 300 in the city of Pensacola (population = 60,600) and 240 in 22 small towns, boroughs, and communities based on the US Postal Service mail delivery in Baldwin County, Alabama (*M* populations = 6,713). In the second experiment, a total of 300 lost letters, i.e., either 25 letters for each of three affiliations for addressees in two sizes of community and for two conditions that manipulated "person-positivity," were distributed in public places in Ohio: 150 letters within the city limits of Lima (population = 45,550) and 150 within the city limits of Ada and Kenton (*M* populations = 6,713). Public places included in the aisles of stores, on frequently used sidewalks next to major thoroughfares, in and around phone booths and automatic (bank) teller machines. The content of the letters in both studies indicated that a $250 contribution was available if someone from one of the respective committees would drop by and pick it up. The name Francis S. Bridges and a residential Pensacola, Florida address was used for the portion of the study conducted in Florida and Alabama and the name Margaret A. Windsor and a Post Office Box address in Lima, Ohio was used for the portion of the study conducted in Ohio. The names of addresses could not be the same in the two field experiments because a valid driver's license is checked by name as a requirement to rent a Post Office Box, and experience has shown that using a name other than the one used to rent the Box may cause delivery problems. Thus, half of the envelopes were addressed to either Francis S. Bridges or Margaret A. Windsor in care of (the) Advocates for Battered and Abused Women (ABAW), Francis S. Bridges or Margaret A. Windsor in care of Advocates for Battered and Abused Men (ABAM), and Francis S. Bridges or Margaret A. Windsor in care of Advocates for Battered and Abused Lesbians (ABAL); the rest of the envelopes simply addressed to these groups. "Francis S. Bridges or Margaret A. Windsor" were used as the name portion of the 'person(s) in care of a group' condition and the Advocates for Battered and Abused Women (ABAW), Advocates for Battered and Abused Men (ABAM), and Advocates for Battered and Abused Lesbians (ABAL) affiliates were used as the 'groups' in both types of addresses.

Research Design and Statistical Analysis

The research design used a hierarchical log linear model (Kinnear & Gray, 2001) to examine both (1) the complex relationships between addressee affiliation (emotive/control), community size (city/small towns), geographic location (Ohio/Florida and Alabama), and return response (returned/not returned), and (2) the structure present in a contingency table. Chi-squared goodness-of-fit testing confirmed the improvements in the fit of the hierarchical log linear model after adding combinations of interactions. Follow-up chi-squared analyses were used to examine (1) the interactions among the variables of addressee affiliation and geographic location on return response to a lost letter. Yates' corrected chi-squared was computed for all 2 x 2 tables. Statistical power for the chi-square analysis used in this study was calculated using *The Program*, by Glantz (1997). All differences were significant at $p < .05$ or beyond.

RESULTS

There was significant interaction for geographic location by return response [X^2 ($N = 840$) = 38.32, $p < .001$, effect size = .21] indicating that the overall returned responses of lost letters differed across geographic location and return response categories regardless of condition. Residents of Ohio (70.0%) had a higher percentage of letters returned than residents of Florida and Alabama (47.6%). The rates in each condition are depicted in Tables 1 and 2.

TABLE 1 Number & Percent of Letters Returned in Ohio as a Function of Affiliation of Addressee, Location, & Object Evoking Attitude

		Condition			
		ABAL[*] (Deviant)	Nondeviant		Total
			ABAW[†]	ABAM[‡]	
City					
Individual Name	*n*	17	20	23	60
Group Name	*n*	20	39	18	61
Both Names	*n*	37	43	41	121
	%	74.0	86.0	82.0	80.7
Town					
Individual Name	*n*	11	14	18	43
Group Name	*n*	17	13	16	46
Both Names	*n*	28	27	34	89
	%	56.0	54.0	68.0	59.3
Total					
Individual Name	*n*	28	34	41	103
Group Name	*n*	37	36	34	107
Both Names	*n*	65	70	75	210
	%	65.0	70.0	75.0	70.0

[*] ABAL = Advocates for Battered and Abused Lesbians
[†] ABAW = Advocates for Battered and Abused Women
[‡] ABAM = Advocates for Battered and Abused Men

There was significant interaction for community size by return response [X^2 (N = 840) = 21.53, $p < .001$, effect size = .16] indicating that the overall returned responses of lost letters differed across community size and return response categories regardless of condition. City residents (63.1%) had a higher percentage of letters returned than residents of small towns (46.9%).

There was a significant interaction for addressee affiliation by return response [X^2 (N = 840) = 23.42, $p < .001$, effect size = .17] indicating that returned responses of lost letters differed between addressee affiliations and return response categories. Letters addressed to Advocates for Battered and Abused Women (ABAW) (62.5%) were returned more often than those letters addressed to Advocates for Battered and Abused Men (ABAM)(60.4%) or Advocates for Battered and Abused Lesbians (ABAL) (43.9%). Since there was significant variation across kinds of

TABLE 2 Number & Percent of Letters Returned in Florida & Alabama as a Function of Affiliation of Addressee, Location, and Object Evoking Attitude

		Condition			
		ABAL[*](Deviant)	Non-deviant		Total
			ABAW[†]	ABAM[‡]	
City					
Individual Name	*n*	20	34	36	90
Group Name	*n*	18	29	26	73
Both Names	*n*	38	63	62	163
	%	38.0	63.0	62.0	54.3
Town					
Individual Name	*n*	16	21	19	56
Group Name	*n*	4	21	13	38
Both Names	*n*	20	42	32	94
	%	25.0	52.5	40.0	39.2
Total					
Individual Name	*n*	36	55	55	146
Group Name	*n*	22	50	39	111
Both Names	*n*	58	105	94	257
	%	32.2	58.3	52.2	47.6

[*] ABAL = Advocates for Battered and Abused Lesbians
[†] ABAW = Advocates for Battered and Abused Women
[‡] ABAM = Advocates for Battered and Abused Men

addressees' affiliations in the overall rates of return, a contingency table was subdivided to compare Advocates for Battered and Abused Women (ABAW) and Advocates for Battered and Abused Men (ABAM) affiliates [X^2 ($N = 560$) = .19, *ns*, effect size = .02]. The returns for these affiliates were pooled and compared to those returns of the Advocates for Battered and Abused Lesbian (ABAL) affiliate. Consequently, addressees' affiliation was significantly associated with overall rates of return [X^2 ($N = 840$) = 22.45, $p < .001$, effect size = .16], with the rates of return being the lowest (61.4% vs. 43.9%) for the homosexual affiliate.

The manipulation of the "person-positivity" condition was significantly associated with different rates of return [X^2 ($N = 840$) = 12.09, $p<.001$, effect size = .12]. There was a significantly larger percentage of lost letters (61.9% vs. 49.8%) associated with returns addressed to 'person in care of a group' (i.e., Francis S. Bridges and Margaret A. Windsor in care of a group) than those addressed to 'group(s)' (i.e., ABAW, ABAM, ABAL).

DISCUSSION

Trying to come up with two fictitious control conditions with different goals, i.e., Advocates for Battered and Abused Women (ABAW) and Advocates for Battered and Abused Men (ABAM), after selecting a suitable experimental condition was a difficult task. Therefore, the addressee affiliation by return response analyses were followed with simple comparisons (initially pooling the ABAW and ABAM conditions) because we could not be sure that public attitudes were not evoked through the use of either one of these affiliates, say for example, in persons angry with their local Shelter for Battered Women.

The log linear results suggest that geographic location, community size, addressee affiliation, and response rate were not independent, and thus, independence does not adequately explain the structure that is inherent in the contingency table. The two-way interaction model (addressee affiliation-return response, geographic location-return response, community size-return response, and geographic location-community size) was adequate to represent the data. Examination of the two-way interaction term of geographic location-community size was the only effect that was found to be non-significant. The finding of a significant difference in response rate for these four variables suggests that the impact of addressee affiliation, geographic location, "person-positivity" and community size appears to account for most of these differences.

The finding of a significant difference associated with return response for the three addressee affiliations appear to support in part, the basic premise of Milgram's (1970) lost letter technique: namely, that the probability of lost letters being returned depends on the social and political attributes of letters' affiliations, i.e., addressees' affiliations. Milgram (1969) concluded "The lost letter technique is not very good for subtle issues, or in connection with issues that do not arouse very strong feelings. It only works for issues in where there is clear-cut polarization, and which arouse a high level of emotional involvement" (p. 264). As to the limitations of the lost letter technique being less likely to make fine distinctions (in our Southern and Northern cities), it seems that our analysis, like that of Waugh's *et al.* (2000) gay marriage analysis, was not too subtle a distinction for the technique. Our findings from mid-size cities, not towns, are consistent with some studies of public attitudes toward homosexuals and other issues of sexual orientation, but not consistent with others. One of these studies (Bridges & Rodriguez, 2000) showed no difference in the rate of returned lost letters within a city, whereas five others (Bridges, 1996, Bridges, Welsh, Graves, & Sonn, 1997; Bridges, Anzalone, & Anzalone, unpublished; Bridges, Meyer, & Scheibe, unpublished; Waugh , Plake, & Rienzi, 2000) reported fewer lost letters returned from a city for any addressee that was clearly associated with a homosexual group or represented other issues of sexual orientation.

The finding of a significant difference associated with return response for residents of the cities and small towns suggests that community size had an impact on return responses. It was expected that as the size of the communities' populations decreased, there would be an increase in the returned responses; however, returned responses from the cities was higher than that of the small towns. The present finding does not agree with Milgram's (1970) hypothesis that people in small towns help others more often than those in cities. However, unlike previous lost letter studies, recent studies tested Milgram's (1970) hypothesis of rural-urban differences by statistically analyzing returns from the different sizes of community. They found that urban communities had *higher* rates of helping (e.g., returned *significantly* more lost letters) than did less populated beach and rural communities (Bridges & Rodriguez 2000; Bridges, Thompson, & Willers, 2000). Thus, our findings suggest the possibility that there may be some community bias in the willingness of people to help some strangers. That is, regardless of the address on a lost letter, a stranger in need is more likely to have a letter returned from cities than from small towns.

The reason(s) why residents of the North were more helpful in returning lost letters than residents of the South is unclear. One possibility is that the researchers who "lost" letters in the South unintentionally made them more difficult to find than the researcher who "lost" them in the North. Even when the emotive ABAL affiliation was removed from the contingency table, there were larger numbers and percentages of lost letters associated with the geographic North than with the Southern locations ($p < .05$).

The "person-positivity" condition was significantly associated with returned responses for the three addressees' affiliations. These findings are consistent with (Simmons & Zumpf (1983), as our letter return rates were significantly higher across all three conditions for letters addressed to "person(s) in care of a group" (i.e., Francis S. Bridges or Margaret A. Windsor in care of a group) rather than to the "group(s)" alone. Perhaps one reason for these findings is because our "person-positivity" conditions were varied enough on the "personhood" dimension to produce a measurable "positivity" bonus in the form of more returned letters for one condition than the other. Another reason might be that our person-positivity conditions both brought to mind a type of emotive or non-ordinary person, of whom evaluators, e.g., letter finders, felt less similar to. This may have resulted in evaluations, i.e., returned responses, being more positive for control conditions than the emotive one.

As expected, returned responses from the small towns were lower for the emotive lesbian addressee; however, even the Advocates for Battered and Abused Lesbian (ABAL) affiliation garnered a return rate of 43.9%. As such, our attitudinal data have, in fact, supported the basic premise of Milgram *et al.*'s lost letter technique: namely that the probability of lost letters being returned depends on the social and political attributes of the addressees' affiliations. The data suggest that further studies are needed for comparisons of responses to lost letters in different community size settings, especially using new or untested social issues (e.g., Advocates for Battered and Abused Gay Men) known to elicit non-normal distributions of public sentiment.

Conclusions from these results should be viewed cautiously because of the following limitations. First, Francis Bridges, the addressee used in the Florida and Alabama portion of the study, may have been perceived by letter finders as a name for a man or a woman because of different spellings (i.e., "i" for males and "e" for females). So it is possible that the higher rate of return in Ohio may be have been due to a gender-specific effect and less of a person-positive one since a lost letter may have been more likely to have been returned if it was addressed to a female (OH) than a male (FL/AL). Despite the potential for confounding the name Francis may have still been a better choice than the name Stephen which is clearly the name of a male and the gender opposite of Margaret. Interestingly, only in Florida and Alabama and not Ohio was there a significant association with returned responses for the individual and group names, with more letters returned with an individual name (146) than without (111) [X^2 (N = 540) = 8.58, $p<.003$, effect size = .13]. Further testing of gender-specific names as male and female addressees in the same geographic location is warranted. Second, any effect on the differential return rates in Ohio and Florida and Alabama towns may ostensibly seem to be because of differences in the population, i.e., a lower pedestrian density in public places in Florida and Alabama may reduce the likelihood of a letter being found than in Ohio. Empirically, we have found that just because there were fewer public places in the 22 small Florida and Alabama towns for people to go, doesn't mean there are fewer people in those places (they are the only public places they have to go). Also, finding a lost letter and helping by carrying it away may be two different things. In an unpublished study, we found that the correlation between the immediate pedestrian density and the number of non-helpful passersby produced a moderately strong positive association. Thus, the greater the number of subjects passing by a lost letter, the less likely anyone one of them would respond to it.

REFERENCES

Abrahamson, M., & Carter, V. J. (1986). Tolerance, urbanism, and region. *American Sociological Reviews, 51*, 287-294.

Berscheid, E., & Walster, E. H. (1978). *Interpersonal attraction*. Reading, MA: Addison Wesley.

Borg, M. J. (1997). The southern subculture of punitiveness? Regional variation in support for capital punishment. *Journal of Research in Crime and Delinquency, 34(1)*, 24-45.

Bridges, F. S. (1996). Altruism toward deviant persons in cities, suburbs, and small towns. *Psychological Reports, 79*, 313-314.

Bridges, F. S., & Coady, N. P. (1996). Urban size differences in incidence of altruistic behavior. *Psychological Reports, 78*, 307-312.

Bridges, F. S., Anzalone, D. A., & Anzalone, F. L. (unpublished). *Teaching affiliation, community size, symbol bias, and person-positivity bias in return of lost letters.*

Bridges, F. S., Meyer, L. G., Scheibe, J. J. (unpublished). *Affiliation and geographic location in return of lost letters.*

Bridges, F. S., & Rodriguez, W. I. (2000). Gay-friendly affiliation, community size, and color of address in return of lost letters. *North American Journal of Psychology, 2(1)*, 39-46.

Bridges, F. S., Ryan, S., Scheibe, J. J. (2000). Affiliation, community size, and person-positivity effect in return of lost letters. *Psychological Reports, 87*, 1003-1010.

Bridges, F. S., Thompson, P. C., & Willers, P. H. (2000). Lost letter returns on five topics from urban, waterfront, and smaller rural communities. *Psychological Reports, 86*, 1097-1100.

Bridges, F. S., Welsh, R. L., Graves, B. S., & Sonn, M. B. (1997). Differences in lost letter and postal card returns from cities and smaller urban communities. *Psychological Reports, 80*, 363-368.

Bridges, F. S., Williamson, C. B., & Scheibe, J. J. (1998). Abortion affiliation, urban size, and geographic bias of responses to lost letters. *Psychological Reports, 83*, 1107-1112.

Byrne, D. (1971). *The attraction paradigm*. New York: Academic Press.

Ellison, C. G., & Musick, M. A. (1993). Southern intolerance: a fundamentalist effect? *Social Forces, 72(2)*, 379-398.

Glantz, S. A. (1997) *Primer of biostatistics: The Program 4/E Version 4.0.* [Software documentation] New York: McGraw-Hill. Pp. 175-180.

Hansson, R. O. & Slade, K. M. (1977). Altruism toward a deviant in city and small town. *Journal of Applied Social Psychology, 7*, 272-279.

Herek, G. M. (1999). AIDS and stigma. *American Behavioral Scientist*, 42(7), 1102-1112.

Herek, G. M., & Capitanio, J. P. (1999). AIDS stigma and sexual prejudice. *American Behavioral Scientist, 42(7)*, 1130-1147.

Hunter, J. D. (1987) *Culture Wars: The struggle to define America*. Basic Books.

Jacoby, J., & Aranoff, D. (1971). Political polling and the lost-letter technique. *Journal of Social Psychology, 83*, 109-212.

Kaplan, M. F. (1976). Measurement and generality of response dispositions in person perception. *Journal of Personality, 44,* 179-194.

Kinnear, P. R., & Gray, C. D. (2001). *SPSS for windows made simple: release 10.* Psychology Press, Taylor & Francis Group.

Middleton, R. (1976). Regional differences in prejudice. *American Sociological Review, 41,* 94-117.

Milgram, S. (1969). Comment on 'A failure to validate the lost letter technique'. *Public Opinion Quarterly, 33,* 263-264.

Milgram, S. (1970). The experience of living in cities. *Science, 167,* 1461-1468.

Milgram, S., Mann, L., & Harter, S. (1965). The lost-letter technique. *Public Opinion Quarterly, 29,* 437-438.

Nunn, C. Z., Crockett, H. J., & Williams, J. A. (1978). *Tolerance for nonconformity.* New York: Jossey-Bass.

Simmons, C. H., & Zumpf, C. (1983). The lost letter technique revisited. *The Journal of Applied Social Psychology, 13,* 510-514.

Sears, D. O. (1983). The person-positivity bias. *Journal of Personality and Social Psychology, 44,* 233-250.

Sears, D. O., Brown, T. L., & Ditto, P. H. (1982). "Humanness", perceived similarity, and the person-positivity bias. (Unpublished manuscript, University of California at Los Angeles).

Sears, D. O., & Whitney, R. E. (1973). Political persuasion. In I. des. Pool, W. Schramm, F. W. Frey, N. Maccoby, & E. B. Parker (Eds.), *Handbook of communication.* Chicago, IL: Rand McNally (pp. 253-289).

Stouffer, S. (1955). *Communism, conformity, and civil liberties.* New York: Wiley.

Tuch, S. A. (1987). Urbanism, region, and tolerance revisited: The case of racial prejudice. *American Sociological Review, 52,* 504-10.

Waugh, I. M., Plake, E. V., & Rienzi, B. M. (2000). Assessing attitudes toward gay marriage among selected Christian groups using the lost letter technique. *Psychological Reports, 86,* 215-218.

Whitehead, G. I. & Metzger, S. C. (1981). Helping behavior in urban and nonurban settings. *The Journal of Social Psychology, 114,* 295-296.

Weiner, M. J. (1975). A nonreactive predictor of a "liquor by the drink" referendum. *The Journal of Social Psychology, 97,* 103-106.

Wicker, A. W. (1969). A failure to validate the lost-letter technique. *Public Opinion Quarterly, 33,* 260-262.

Wilson, T. C. (1986). Interregional migration and racial attitudes. *Social Forces, 65,* 177-186.

Filicide-suicide: In Search of Meaning

Perry L. Collins
Wayland Baptist University

Michael F. Shaughnessy
Eastern New Mexico University

Loretta Bradley
Texas Tech University

Kaylene Brown
Texas Tech University

Each year a few parents murder their children and then commit suicide. A review of the literature suggests two primary motives for a parent to commit such acts. The first motive involves actively attempting to harm the child or children, which includes both accidental death as a result of physical abuse or actual murder as the primary intent. In this *murder motive*, the parent may then commit suicide after feeling a deep sense of despair and hopelessness following an appraisal of their behavior toward their children. The second motive, or *suicide motive*, involves the parent premeditating suicide and choosing to take the lives of their offspring as well. In this "mercy killing" manner the intent is to relieve the current or future suffering or pain of their offspring. This article explores various factors in parental filicide-suicide and suggests what counselors should do to prevent such tragedies.

What do Mark Heath, Barbara Wyrzykowski and Ronald Jonker all have in common? To most readers, these are just names. However, these Australians have recently come to the attention of the world in that they all committed suicide and took their children with them. Heath gassed himself and his four children, Wyrzykowski took her own life and the lives of her five children, and Jonker gassed himself and his three children on the northern outskirts of Perth. Australian newspapers placed the pictures of these innocent children on the front pages.

These occurrences have caused untold turmoil to the survivors of the families and have left social workers, psychologists, and therapists puzzled and concerned. While homicide-suicides have not received much research attention, this is not a new phenomenon. Homicide-suicides date at least as far back as the Ming dynasty (Rin, 1975) and even earlier

---

Author info: Correspondence should be sent to: Dr. Michael Shaughnessy, Psychology Dept, Eastern New Mexico University, Portales, NM 88130.

North American Journal of Psychology, 2001, Vol. 3, No. 2, 277-292.
© NAJP

mentions can be found in Greek mythology. Sorenson and Peterson (1994) found that over the past several decades homicide was one of the leading causes of death in children in the United States. Furthermore, their research indicates that the rate of child homicide continues to be on the rise. In Finland, 2% of all deaths of children under the age of 15 years resulted from homicide in 1995, and in 70%-80% of child homicides in recent years, the perpetrator has been a family member (Statistics Finland, 1996). While the intent of this article is not to explore child homicide, it is important to note that oftentimes one or both parents of these children are the ones responsible for their homicide. Some of these parents choose to take their own lives following the act of filicide. The purpose of this article is to explore the phenomenon of filicide-suicide and to consider ways in which counselors might help to address this problem. The topic of filicide-suicide within the broader context of murder-suicide will be discussed in terms of its prevalence worldwide, current research findings, the various methods of murder in murder-suicides, different motives for committing murder-suicide, and common factors of the murdering parent as well as the child victim. This article will conclude with a discussion of implications and considerations for counselors.

Murder-suicides occur worldwide

In an existential crisis, to take one's own life is, of course, a reflection of the lack of meaning in one's own life. To take the lives of one's own children is usually a senseless act. Certainly, these incidents do not take place in Australia alone; they can and do occur in other countries. A brief article in *Maclean's* on September 6th, 1999 read as follows:

> A 41-year-old Toronto man jumped to his death in front of a speeding subway train, dragging his three-year-old son with him. Jeyabalan Balasingam, who immigrated to Canada from Sri Lanka in 1993, had been treated for depression for several years, not uncommon for new immigrants who leave family networks behind (p. 21).

A more recent article in *Maclean's* on July 17, 2000 described a similar event:

> One of the worst mass domestic killings in Canadian history occurred in Kitchener, Ont. Four small children were murdered by their father, Bill Luft, 42, who also stabbed his wife to death and shot himself fatally. Authorities said Luft had recently sought help for psychiatric problems (p. 23).

While the two incidents described above occurred in Canada, filicide-suicides take place in the United States as well. The following *Maclean's* excerpt dated March 31, 1997 read:

> Huge debts led a self-made millionaire from Canada to kill his wife, eight-year-old son and then himself in West Palm Beach, Fla. Former Toronto developer Robert McIntosh, 72, shot his family with a .22-calibre handgun. A five-page suicide note said the parents did not want to burden anyone with raising their son and they felt he would be scarred by their deaths. It also said they attempted murder-suicide by carbon monoxide poisoning in January, but the child talked them out of it midway through (p. 41).

Somander and Rammer (1991) studied intra-familial and extra familial child homicide in Sweden for the years 1971-1980. They cited research conducted by Grunfeld and Steen (1984) which indicated that "attempted or successful suicide is a relatively common aftermath of child homicide." (p. 45). Somander and Rammer (1991) defined homicide-suicide as the killing of a child within the context of the suicide or suicide attempt of the perpetrator.

Iga (1996) wrote of the cultural aspects of Japanese *oyako shinju* (parent-child suicide) that occurred in the United States. Iga (1996) found that traditionally *oyako shinju* seemed to occur due to actual or expected financial difficulty, which often occurred after divorce. According to Berman (1979), infanticide-suicide (which is a form of filicide-suicide that involves victims under the age of one) appears markedly less often in the United States than in cultures such as Japan. Haapasalo and Petaja (1999) studied mothers in Finland who killed or attempted to kill their children and related it to the mothers' life circumstances and childhood abuse. Furthermore, Buteau, Lesage, and Kiely (1993) investigated homicide followed by suicide in Quebec, Canada. While their study considered all murder-suicides and not just those involving parents and their children, they do acknowledge that homicide-suicide "often occurs in a family context and may involve children as victims." (p. 552).

Research on murder-suicides

Murder followed by suicide is relatively rare. In the United States, the rate is about 2.2 per million population (Palmer & Humphrey, 1980), and many of these are not parental filicide-suicide. Some may not involve family members at all while others may include a perpetrator who kills their spouse or partner and then commits suicide. West (1967) found a higher incidence of prior suicide attempts among those adults involved in murder-suicides when compared to a matched suicide only group. In a study of murder-suicides conducted in the United States, Berman (1979) found that most of the perpetrators of murder-suicide used a firearm and shot their victim in the head before later killing themselves. Forty percent of the perpetrators in his study were under the influence of alcohol and fifty percent had prior arrests for criminal behavior. When comparing his study with similar studies conducted by Selkin (1976) and Wolfgang (1968), he found that an overwhelming majority of cases involved only one other victim and used a violent method for carrying out the act. Most of the offenders in these three studies were male with a mean age ranging from 38 to 43 years old. The ethnicity of the offenders in these studies (which were conducted in the United States) tended to be Caucasian.

West (1967) and Dorpat (1966) both argued that a majority of perpetrators of murder-suicide are psychotic. While there is little follow-up research to substantiate these claims, Berman (1979) believes that many of these offenders had an unhealthy relationship with the individuals who later became the victims of the murder-suicide act. He further stated that these relationships were formed and sustained in a manner in which the two "depended on and abused each other for affection, displaying chronic love-hate patterns, omnipotent expectations, narcissistic demands, and definitions of love rigidly posed." (p. 21).

Berman's study did not focus exclusively on filicide-suicide. However, some of the same factors may be relevant. Selkin (1971) reported that the fantasy of rescue, which involves freeing survivors from the pathology of the family system, may be a dominant motive in many murder-suicides.

Berman (1979) posited an intriguing motivation for why murder-suicide occurs. He argued that a dominant theme in murder-suicides is that of dependent-protective. He concluded that, "rather than with ambivalence and rage, the homicide victim is viewed as dependent upon and needing of the care-taking functions of the perpetrator." (p. 22). He also indicated that when this protective, supportive role is threatened through loss of money, health, etc., feelings of hopelessness and frustration might lead to this dependent-protective motivation for the murder-suicide. Even though in many cases of murder-suicide the primary motivation is murder, it does appear that suicide can also be a primary motivation. When this occurs, the homicide has been viewed as a "merciful extension of the fused ego state" between the perpetrator and the victim (Cavan, 1965, p. 22). According to Leonard (1967), the dominant theme when suicide is the primary motive is the inadequacy of differentiation between self and object. In other words, according to neo-analytic thought, differentiation is the process by which an individual learns to chart his or her own direction in life. A parent may have some unresolved childhood attachment issues with their mother and may seek to resolve such issues by becoming so enmeshed with their own child that it creates a state of fusion. Thus, the child becomes merely an extension of the parent's self.

Regardless of the reasons proposed, the fact remains that the occurrences of filicide-suicide should serve to alert mental health professionals. Counselors should understand that there is something wrong with some individual value systems, and that people in the helping professions have not always successfully assisted others in finding meaning in their lives.

Method of murder in filicide

According to Haapasalo and Petaja (1999), mothers who killed or attempted to kill their child or children used a variety of methods, including drowning, drugging, strangulation, suffocation, neglect, physical abuse, and stabbing. In their examination of 56 cases, they also found other means of murder. These included the following: throwing their child in front of a moving vehicle, putting pet shampoo into an infant's bottle for the purpose of poisoning the baby, striking a child with a blunt object (a rock), and throwing a child into a wall. Haapasalo and Petaja (1999) even found two cases in which mothers used multiple methods of lethality simultaneously. In one instance a mother poisoned and suffocated her child and in another case a mother stabbed and strangled her child. In younger children, drowning in a bathtub has been one of the most common methods (Haapasalo & Petaja, 1999). Cooper and Eaves (1996) investigated a case in which a mother killed her daughter and herself by carbon monoxide poisoning. Buteau, Lesage, and Kiely (1993) reported that firearms were used in 56% of the homicide-suicides that they investigated. According to an article in *Maclean's* dated January 22, 1996, James Huang, a 40-year-old Taiwanese real estate developer, killed four members of his family and then committed suicide using two rifles and a shotgun. Handguns and other types of firearms have also been reported as methods of murder-suicide (Morton, Runyan, Moracco, & Butts, 1998).

From a review of the literature, it does appear that drowning such as in the bathtub seems to be a more common method of murder in younger children although many other means have been used as mentioned above. Furthermore, like suicide, men seem to use more violent and lethal means than women when acting as perpetrator's in these murder-suicides. They often use rifles, shotguns, handguns or other types of firearms as well as knives in carrying out these acts. While women have been known to shoot or stab their victim during a murder-suicide, they more often will resort to less violent measures like poisoning, drowning, and neglect.

Motives for murder-suicides

From a review of the literature, there appears to be two primary motives in a majority of these murder-suicides that involve one or both parents and one or more children. We refer to these two motives as the *murder motive* and the *suicide motive*.

Murder motive. In the first motive, the primary intent of the parent is to inflict harm on their child. In many cases of child abuse that result in death, there are no premeditated thoughts of actually killing their son or daughter but the child may "accidentally" die during the course of the physical abuse. In other cases that involve the *murder motive*, the parent actually intended to commit murder. According to Resnick (1969), this can occur because of acute psychosis, acts related to an unwanted child, and spouse revenge. Scott (1973) found similar reasons for parental filicide. When exclusively considering mothers who kill their children, d'Orban (1979) found six different types which include the following: battering mothers, mentally ill mothers, neo-naticides, retaliating mothers, women who killed unwanted children, and mercy killing. According to Haapasalo and Petaja (1999), revenge is a rare motive for child killing. Crimmins, Langley, Brownstein, and Spunt (1997) referred to the revengeful motive as the "Medea syndrome," after the Greek tragedy of Medea, a jealous, vengeful wife who took her revenge on her husband by murdering their two children.

When the child dies either of accidental death secondary to physical abuse or when murder is the primary intent, the parent who committed the act may feel devastated and overcome with guilt. The parent then may subsequently make an attempt on their own life as a means of escape from the tremendous pressure created by their actions. There may be incidents reported where an infant dies of SIDS and the mother later commits suicide following such a traumatic event. However, according to Emery (1993), maternal filicide may occasionally remain undetected. Emery (1993) found that up to 20% of all SIDS deaths might be attributed to something other than natural causes. Child death due to factitious disorder by proxy may be difficult to assess as well. With factitious disorder by proxy there is a deliberate production of feigning of physical or psychological signs or symptoms in another person who is under the individual's care. Typically the victim is a young child and the perpetrator is the child's mother. The motivation for the mother's behavior seems to be a psychological need to assume the sick role by proxy. These mothers seem to find it very rewarding to nurture and take care of their "sick" children. Unfortunately there may be times when a mother with factitious disorder by proxy may do things to actually hurt the child or make the child sick in an effort to take the role of caregiver. On these occasions there is always a risk of accidentally killing the child in the process. Up to 85% of factitious disorder by proxy (also referred to as Munchhausen syndrome by proxy) involves the mother as the perpetrator (McClure, Davis, Meadow, & Sibert, 1996).

When parents actually murder their children, they may do so for a variety of reasons. Some individuals who murder their children, then plan on turning the gun on themselves, may at the last minute change their minds, or be overcome with grief, remorse or be overwhelmed with what they have done and begin to re-evaluate and reassess their actions. They may then be immobilized with the reality of what they have done and feel that they deserve punishment. There may also be outside intervention on the part of police or neighbors.

Parents may kill their children due to paranoid delusions or hallucinations of auditory control. In this scenario the active phase of schizophrenia, schizotypal personality disorder, delusional disorder, brief psychotic disorder, or shared psychotic disorder may account for their actions. Haapasalo and Petaja (1999) argue that a theoretical model of the psychological processes underlying parental homicidal acts against a child is lacking in the literature. They posit that, from a psychological viewpoint, "there appears to be no compelling reason to treat them [parental homicidal behaviors] as separate phenomena [from homicides that do not involve a parent]." (p. 220).

Suicide motive. In many cases of murder-suicide involving parents and children, murder is not the primary motivation. In these instances, the primary motivation for the murder-suicide is the suicide of the parent. Some parents feel so depressed and hopeless that suicide becomes perceived as their only option. The parent may then worry about the welfare of his or her children and may believe that taking the child's life as well will save the child from current or future suffering. Men continue to have higher rates of suicide than women (Stack, 2000) and this may explain why more fathers than mothers engage in filicide-suicide.

In 41 of 42 nations surveyed (Kuwait being the exception), men had higher rates of suicide than women (Travis, 1990). Stack (2000) described many reasons for men having higher rates of suicide which include the following: 1) males abuse alcohol more frequently, which can reduce their inhibitions (Canetto, 1992); 2) women are more religious, which reduces the likelihood that suicide will be considered as a viable option in the facing of mounting stress (Stack, 1983); 3) women have stronger negative attitudes towards suicide attempts (Canetto, 1992); 4) women have more flexible coping skills which provide them with options to consider other than suicide (Girard, 1993); 5) women are more likely to recognize and less likely to deny the warning signs of suicide, such as depression, which may prompt them to seek professional help (Canetto, 1992); 6) women are more likely to seek professional help, which can lower their risk of suicide (Canetto, 1992); 7) women have more extensive social support systems than men (Pescosolido & Wright, 1990); 8) men may be more impulsive even when not under the influence of drugs or alcohol (Stillion, 1984); 9) men have more access to lethal means of successfully committing suicide, such as firearms (Kushner, 1985); and 10) failure in the primary adult male role (economic success) is more visible and obvious than failure in the primary adult female role (success in relationships) (Girard, 1993). This can lead to greater feelings of desperation on the part of men compared to women.

Resnick (1969) studied 88 cases of maternal filicide and classified them according to their apparent motives. He found that in most cases the filicide was altruistic. The mother either committed suicide and took her child with her or the filicide was committed to relieve real or imagined suffering by the child. When suicide is the primary motivation for the murder-suicide, many of the same psychosocial factors that lead to suicidal behavior in non-homicidal parents are also relevant in explaining the suicidal behavior of parents who commit filicide-suicide (Cooper & Eaves, 1996; Fishbain, Rao, & Aldrich, 1985). For instance, a parent's poor health may be a factor in their consideration of committing suicide. Another factor may be depression or other forms of mental illness.

Cooper and Eaves (1996) studied two female filicide-suicide perpetrators. They described the first case as a "psychotic break" in which the female murdered her husband and two sons due to mental illness. In the second case, in which a woman killed herself and her daughter by carbon monoxide poisoning, Cooper and Eaves (1996) mentioned that the woman had been consulting a psychiatrist for depression. This woman had left suicide notes that indicated extreme sensitivity to perceived slights from members of her family. The perpetrator indicated that her motive was revenge against her mother and ex-boyfriend. Thus, both of these cases seem to fit Daly and Wilson's (1988) "insanity" explanation of murder-suicide.

Common variables that lead parents to murder-filicide

From the literature examined, there is no model that definitively explains both murder-suicide which did not involve family members and filicide-suicide. However, several common variables were revealed in the murder-suiciding parent. These commonalties ranged from the influence of the parent's own childhood abuse (Haapasalo & Petaja, 1999) to altruistic reasons citing the "well-being of the child as a primary concern" (Somander & Rammer, 1991, p. 49) to an intense love for the children which prompted a "mercy killing" (Iga, 1996). While men tend to kill their children more often than women, mothers are more frequently perpetrators than fathers during the child's first year of life which is referred to as parental neo-naticide (Fornes, Druilhe, & Lecomte, 1995). This may result from the fact that mothers tend to spend significantly more time with their infant children than fathers. Increased exposure may exacerbate the stress experienced by the mother in caring for her infant. Mothers are also more likely to treat their children as extensions of themselves as compared to fathers, who are less likely to treat children as inseparable parts of themselves. This may increase the mother's risk for murdering her child and then taking her own life (Myers, 1970). Crittenden and Craig (1990) studied 171 cases of children under 13 years killed in one county from 1956 to 1986 and found that mothers were perpetrators in 86% of neonatal deaths and in 39% of infant deaths.

According to McKee and Shea (1998), a majority of mothers who kill their children have been found to be single, young, and living in difficult circumstances. These mothers appear to be immature in many respects and their homicidal acts may be related to severe psychological stress in their daily lives (Haapasalo & Petaja, 1999). This immaturity "may be manifested in their inability to meet socially and emotionally the requirements and responsibilities of adult life and the role of a mother." (p. 220).

Another at-risk group, according to Haapasalo and Petaja (1999), are older, married mothers who are experiencing marital difficulties. Homicidal behavior in these cases is usually precipitated by the child's behavior. The mother responds with aggression. For this at-risk group the child often becomes a vicarious object of aggression felt toward a husband or some other person (Silverman & Kennedy, 1988).

With regard to ethnicity as a variable in filicide-suicide, the motive for the filicide-suicide becomes a factor. If the primary motive is suicide, then the perpetrating parent is most likely to be a Caucasian. There is evidence in the literature that the suicide rate in the United States among Caucasians is double the suicide rate of African Americans (Lester, 1990). If, however, the primary motive in the filicide-suicide is murder, then African American parents may be at an increased risk of engaging in filicide-suicide.

When confronted with frustration, African Americans are more apt to blame society or others and to externalize aggression in such forms as homicide. In contrast, Caucasians are less likely to attribute their various social and economic failures to discrimination. As such, when Caucasians are confronted with frustration, they are more apt than African Americans to attribute the frustration to their own inadequacies. Aggression, then, is more likely to be turned against oneself in such forms as suicide (Stack, 2000, p. 147).

As mentioned earlier, depression and other psychiatric disorders may account for some of the murder-suicide behavior. According to

Haapasalo and Petaja (1999), "exposure to acute stressors and traumatic events with the subsequent posttraumatic stress symptoms might account for some of the mental disorders observed in homicidal mothers." (p. 220). Murray and Carothers (1990) found postpartum depression issues in approximately 10% of women, and this might account for some of the homicidal behavior described above. Finally, a parent's own history of child abuse, or traumatization may play a role in filicide-suicide behavior (Haapasalo & Petaja, 1999).

Crimmins et al. (1997) found that most women convicted of killing their child reported having experienced or witnessed physical and/or sexual abuse in childhood. And while not all abused parents continue the cycle of violence (Kaufman & Zigler, 1993), many women who had killed their child reported childhood experiences of "severe and continuous verbal abuse, parental alcoholism, parental mental health problems, and the absence of the mother due to illness or death." (p. 222). It should be noted, however, that these women who participated in the study conducted by Crimmins et al. (1997) had not committed suicide following the murder of their child. Therefore, it is uncertain as to what extent these variables apply to mothers engaged in murder-suicide acts.

Stress can play a significant role in murder-suicide behavior by parents. Fathers may be frustrated due to financial difficulties, fatigue, conflicts with family members, poor housing conditions, work-related stress, and child-related misbehaviors. Mothers may experience some of the frustrations that the fathers experience as well. However, mothers may also struggle with issues such as the physiological changes and emotional lability of pregnancy, their partner's alcohol abuse, the stress involved in miscarriage, a lack of social relationships, their partner's absence from the home for extended periods, and the stress involved in carrying out various care-giving tasks (Haapasalo & Petaja, 1999).

There is little in the current body of literature that suggests a causal role for the media in filicide-suicide. However, there is evidence that, given the high rate of marital disruption in many countries, stories concerning suicides or murder-suicides of the divorced or people with marital trouble might serve as a point of mass identification for copycat suicides or murder-suicides. Stack (2000) found that stories about suicide victims with marital trouble were five times more closely linked to copycat suicides than all other publicized stories.

Finally, just as the influence of culturally important temporal factors and ceremonial occasions can affect suicide, these same factors may also play a role in filicide-suicide. According to Stack (1995), holidays, weekends, and springtime tend to be associated with rising expectations for fulfillment. "If these expectations are not met, it seems plausible that suicidal behavior might increase." (p. 153). Gabennesch (1988) argues that holidays may be filled with "broken promises" for some family members.

> Holidays raise expectations for individual fulfillment. When these expectations are not fulfilled by the holiday, suicide risk increases. Furthermore, frustrations involving such broken promises might also explain any rise in suicidal behavior after a weekend ("blue Monday") and in the freshness of springtime (p. 129).

Chew and McCleary (1995) found evidence that the degree of seasonality in suicide in their study of 28 nations was related to modernization. The greater the percentage of the labor force in agriculture, the greater the seasonality of suicide. This finding is consistent with a classic Durkheimian perspective on seasonality in suicide in that it is a consequence of seasonality in social activity (e.g., farming). While this evidence speaks more to suicidal behavior rather than filicide-suicide, seasonality may influence filicide-suicide behavior. If factors of suicide are associated with filicide-suicide, then socioeconomic status and unemployment may play a role as well.

Common features of the child victim

Studies suggest a bi-modal distribution. In other words, children are most often killed by their parents either during the first years of life or in adolescence (Fornes, Druilhe, & Lecomte, 1995; Marks & Kumar, 1996). Many of these children have had a history of prolonged child abuse and neglect (Crittenden & Craig, 1990). Many of them were unplanned or unwanted as well (Resnick, 1969). Although the child victims tend to be in good health (Haapasalo & Petaja, 1999), there are instances where the child has medical problems or cognitive disabilities and a parent considers a "mercy killing" to protect their child from prolonged suffering or perceived future struggles.

Conclusion

There no doubt have been talk show hosts discussing not only filicide-suicide but also similarly unexplainable events involving murder-suicide such as those that occurred in Littleton, Colorado. There will be questions such as "Could this have been prevented?" While at present there seem to be no definitive answers, these events may have been prevented if appropriate identification and intervention had occurred. Identification and intervention must have as their foundation a clear understanding of the problem. There are still many unknowns in the study of filicide-suicide, and it is hoped that the present article will generate more interest and research on the topic.

What should counselors and therapists do?

From the perspective of an existentialist counselor or therapist a significant aspect of our work is to help restore meaning in the lives of those whom we serve. Many counselees struggle with existential crises in their daily lives. Part of our role may not just be to polish our ability to identify and intervene with people in these situations but also to make the public more aware of things that can be done to minimize such occurrences. Does simply posting the pictures of child victims on the front page of the newspapers help in addressing this problem? It may serve the purpose of awareness, but as the media move on to other significant events in the world, the problem remains. In Frankl's book entitled *Man's Search for Meaning,* which was revised and updated in 1984, he cites a study conducted in France in which 89% of people polled admitted that "man needs something for which to live" (p. 121-122). By having something to live for, people may re-evaluate their thoughts of harming themselves or others.

Counselors and therapists (who work specifically towards helping counselees find a sense of meaning in love as well as pain and suffering) should help people to find significance in their existence and should assist them in finding meaning in their family (including their children). Frankl (1984) wrote that "existential frustration can result in neurosis" (p. 123). As mentioned earlier, one of the factors that may play a role in filicide-suicide is mental illness.

Helping others in reducing or coming to terms with their "existential frustration" may reduce their urges to resort to drastic measures such as murder or suicide. Finding meaning in life can reduce this frustration. In the words of Nietzsche, "He who has a *why* to live for can bear almost any *how*."

Frankl (1984) alludes to a possible explanation for holidays and weekends being a time of higher incidence of suicide, and possibly filicide-suicide. He refers to "Sunday neurosis" as a "kind of depression which afflicts people who become aware of the lack of content in their lives when the rush of the busy week is over and the void within themselves becomes manifest." (p. 129). When others come to understand themselves and find purpose in daily life, this existential vacuum becomes less of a factor to them. A categorical imperative of logotherapy which should be imparted to counselees is as follows: "Live as if you were living already for the second time and as if you had acted the first time as wrongly as you are about to act now!" (Frankl, 1984, p. 131-132).

Frankl believed that people can discover meaning in their lives: 1) by creating work or doing things for others, 2) by experiencing something or encountering someone, and 3) by the attitude they take toward unavoidable suffering (p. 133). Counselors and therapists can apply these principles by encouraging clients to look for ways to love and serve others and to appreciate the beauty in their surroundings.

Counselors must also help people find meaning in the parenting of their children. This can be done in a variety of ways, including starting parenting classes and forming parent support groups. Therapists must cautiously take a longer, more critical, clinical look at the potential for violence that their clients may possess. School personnel and school board members must rethink their beliefs that "it can't or won't happen here in our schools." While there have not been documented cases of filicide-suicide in school, mental health counselors should be prepared for this eventuality. With the amount of stress in the field of teaching it is not inconceivable that this event could occur.

This invincibility fallacy which is commonly observed in the egocentrism of adolescents has no place in the school system. As therapists, we recommend taking a proactive stance and working to help people find meaning in their lives. If other counselors and therapists are willing to do this the tragedy of filicide-suicide will occur less frequently.

REFERENCES

Berman, A. L. (1979). Dyadic death: Murder-suicide. *Suicide & Life-Threatening Behavior, 9* (1), 15-23.

Buteau, J., Lesage, A. D., & Kiely, M. C. (1993). Homicide followed by suicide: A Quebec case series, 1988-1990. *Canadian Journal of Psychiatry, 38*, 552-556.

Canetto, S. S. (1992). Gender and suicide in the elderly. *Suicide and Life-Threatening Behavior, 22*, 80-97.

Cavan, R. (1965). *Suicide*. New York: Russell & Russell.

Chew, K., & McCleary, R. (1995). The spring peak in suicide: A cross-national analysis. *Social Science in Medicine, 40*, 223-230.

Cooper, M., & Eaves, D. (1996). Suicide following homicide in the family. *Violence and Victims, 11* (2), 99-112.

Crimmins, S., Langley, S., Brownstein, H. H., & Spunt, B. J. (1997). Convicted women who have killed children: A self-psychology perspective. *Journal of Interpersonal Violence, 12*, 49-69.

Crittenden, P. M., & Craig, S. E. (1990). Developmental trends in the nature of child homicide. *Journal of Interpersonal Violence, 5*, 202-216.

Daly, M., & Wilson, M. (1988). *Homicide*. New York: Aldine de Gruyter.

Dorpat, T. L. (1966). Suicide in murderers. *Psychiatry Digest, 27*, 51-55.

Emery, J. L. (1993). Child abuse, sudden infant death syndrome, and unexpected infant death. *American Journal of Diseases in Children, 147*, 1097-1100.

Fishbain, D. A., Rao, V. J., & Aldrich, T. E. (1985). Female homicide-suicide perpetrators: A controlled study. *Journal of Forensic Science, 30* (4), 1148-1156.

Frankl, V. (1984). *Man's Search for Meaning (revised and updated.* New York: Washington Square Press.

Fornes, P., Druilhe, L., & Lecomte, D. (1995). Childhood homicide in Paris, 1990-1993: A report of 81 cases. *Journal of Forensic Sciences, 40,* 201-204.

Gabennesch, H. (1988). When promises fail: A theory of temporal fluctuations in suicide. *Social Forces , 67,* 129-145.

Girard, C. (1993). Age, gender and suicide: A cross-national analysis. *American Sociological Review, 58,* 553-574.

Grunfeld, B., & Steen, J. (1984). Fatal barnemishandling: Barnedrap I Norge (Fatal child abuse in Norway, in Norwegian). *Tidskrift for Norsk Laegeforening, 104,* 289-292.

Haapasalo, J., & Petaja, S. (1999). Mothers who killed or attempted to kill their child: Life circumstances, childhood abuse, and types of killing. *Violence and Victims, 14(*3), 219-239.

Iga, M. (1996). Cultural aspects of suicide: The case of Japanese oyako shinju (parent suicide). *Archives of Suicide Research, 2,* 87-102.

Kaufman, J., & Zigler, E. (1993). The intergenerational transmission of abuse is overstated. In R. J. Gelles & D. R. Loseke (Eds.),*Current controversies on family violence* (pp. 209-221). Newbury Park, CA: Sage Publications.

Kushner, H. (1985). Women and suicide in historical perspective. *Signs, 10,* 537-552.

Leonard, C. V. (1967). *Understanding and preventing suicide.* Springfield, Illinois: Charles C. Thomas.

Lester, D. (1990). Mortality from suicide and homicide for African Americans in the USA: A regional analysis. *Omega, 22,* 219-226.

Marks, M. N., & Kumar, R. (1996). Infanticide in Scotland. *Medicine, Science and Law, 36,* 299-305.

McClure, R. J., Davis, P. M., Meadow, S. R., & Sibert, J. R. (1996). Epidemiology of Munchhausen syndrome by proxy, non-accidental poisoning, and non-accidental suffocation. *Archives of Diseases in Children, 75,* 57-61.

McKee, G. R., & Shea, S. J. (1998). Maternal filicide: A cross-national comparison. *Journal of Clinical Psychology, 54,* 679-687.

Morton, E., Runyan, C. W., Moracco, K. E., & Butts, J. (1998). Partner homicide-suicide involving female homicide victims: A population based study in North Carolina, 1988-1992. *Violence and Victims, 13* (2), 91-106.

Murray, L., & Carothers, A. D. (1990). The validation of the Edinburgh post-natal depression scale on a community sample. *British Journal of Psychiatry, 157,* 288-290.

Myers, S. A. (1970). Maternal filicide. *American Journal of Diseases of Children, 120,* 534-536.

d'Orban, P. T. (1979). Women who kill their children. *British Journal of Psychiatry, 134,* 560-571.

Palmer, S., & Humphrey, J. (1980). Criminal homicide followed by offender's suicide. *Suicide and Life Threatening Behavior, 10* (2), 106-118.

Pescosolido, B., & Wright, E. (1990). Suicide and the role of the family over the life course. *Family Perspective, 24,* 41-60.

Resnick, P. J. (1969). Child murder by parents: A psychiatric review of filicide. *American Journal of Psychiatry, 126*, 73-82.

Rin, H. (1975). Suicide in Taiwan. In N.I. Farberow (Ed.), *Suicide in different cultures*. Baltimore: University Park Press.

Scott, P. D. (1973). Parents who kill their children. *Medicine, Science and the Law, 13*, 120-126.

Selkin, J. (1976). Rescue fantasies in homicide-suicide. *Suicide and Life-Threatening Behavior, 6*, 79-85.

Silverman, R. A., & Kennedy, L. W. (1988). Women who kill their children. *Violence and Victims, 3*, 113-127.

Somander, L. K. H., & Rammer, L. M. (1991). Intra- and extrafamilial child homicide in Sweden 1971-1980. *Child Abuse and Neglect, 15*, 45-55.

Stack, S. (1983). The effect of religious commitment on suicide: A cross-national analysis. *Journal of Health and Social Behavior, 24*, 362-374.

Stack, S. (1995). The effect of holidays on race specific suicide and homicide rates. *Sociological Focus, 28*, 313-328.

Stack, S. (2000). Suicide: A 15-year review of the sociological literature (Part 1: Cultural and economic factors). *Suicide and Life-Threatening Behavior , 30* (2), 145-162.

Statistics Finland. (1996). *Causes of death*. Health 5. Helsinki: Author.

Stillion, J. (1984). Perspectives on the sex differential in death. *Death Education. 8*, 237-256.

Travis, R. (1990). Suicide in cross-cultural perspective. *International Journal of Comparative Sociology , 31*, 237-248.

West, D. J. (1967). *Murder followed by suicide*. Cambridge: Harvard University Press.

Wolfgang, M. E. (1958). An analysis of homicide-suicide. *Journal of Clinical & Experimental Psychopathology, 19*, 208-218.

Book Review: *Stop Smoking and Chewing Tobacco for Life Changes*

John Karkut
Bloomington, IN

Nicotine addiction is a complex problem. Like a spider web, it ties together one's beliefs about one's self, maladaptive behavior patterns, preoccupations, emotional issues, one's worldview -- in short, every aspect of one's existence into an identity which depends upon smoking to maintain its equilibrium. The task of freeing an individual from nicotine dependence is therefore a complex one.

David and Carole Johnson's book "Stop Smoking and Chewing Tobacco for Life Changes" tackles the problem from many angles. It addresses not only the task of achieving cessation from nicotine use but also provides tools for preventing relapses. Its language is colorful, its techniques imaginative. Yet, unlike other smoking-cessation books which I have read, this book's wide array of interventions are presented in a writing style so casual and a tone so warm and often comical that as one reads it, the book quickly becomes a warm, supportive friend to the individual who struggles, with nerves frayed, to put out that last cigarette or spit out that last chew. The collection of nicotine-cessation techniques presented in the book is truly awesome and includes over 120 exercises in cognitive-behavioral strategies, affirmations, hypnosis, aversion, relaxation techniques, and a host of other strategies targeted to assist the reader in becoming and staying nicotine-free. The homework exercises are not without pain. The aversion procedures, for example, are truly aversive even to read (as they should be). But then, if they were not, how can they be called aversion procedures. The authors' own experience with nicotine cessation shows up in various comforting ways. For example, from time to time, there are proverbs, wise old sayings, and just plain good advice about how to "hang in there" when the going gets tough. The book includes a wealth of resources, e.g. Internet sites, national organizations, readings and other useful materials to assist in the cessation and maintenance process. Techniques for relapse prevention are welcome part of this book. I am all too well acquainted with a number of programs that have not addressed the issue of relapse adequately. In summary, *Stop Smoking and Chewing Tobacco for Life Changes* assists

---

Author info: Correspondence should be sent to: Dr. John Karkut, Box 1396, Bloomington, IN 47402.

North American Journal of Psychology, 2001, Vol. 3, No. 2, 293-294.
© NAJP

individuals in achieving and maintaining freedom from nicotine use while at the same time making the necessary life changes which alter the need for nicotine use. I commend the authors for this comprehensive, self-paced, and imaginative program.

The Role of Religion and Responses to the Y2K Macro-crisis

Larry Powell
Mark Hickson, III
William R. Self
Jean Bodon
University of Alabama at Birmingham

This study examined the relationship between religious activity and responses to the Y2K crisis. A telephone survey of 497 Alabama residents measured their frequency of church attendance and questioned them about (1) their knowledge of the term "Y2K," (2) their perception of the seriousness of the problem, (3) how much information about Y2K that they noticed in the media, (4) how many interpersonal discussions they had on the topic, and (5) whether they made preparations in readiness for Y2K. The results indicated that religious activity was associated with perceptions of the amount of media coverage of Y2K, frequency of interpersonal discussion on the topic, and level of preparations for the possible crisis. Frequent church-goers were more likely to have taken preparations against a Y2K problem, reported seeing more information about Y2K in the media, and engaged in more interpersonal discussions on the topic than did non-churchgoers. No significant difference was identified for either knowledge level or perception of the salience of the issues.

A number of previous research studies have examined how information about critical events is perceived (Gantz, 1983; Greenberg, 1964; Kraus, Davis, Lang & Lang, 1976; Mendelsohn, 1964; Riffe & Stovall, 1989; Steinfatt, Gantz, Siebold & Miller, 1973; Weaver-Lariscy, Sweeney & Steinfatt, 1984). More recently, researchers have looked at public responses to the Y2K macro-crisis (Hickson, Powell, & Bodon, 2001), i.e., concerns over the possibility of computer disruptions after the end of the year 1999. The Y2K "bug" was perceived by some as a "millennial panic" that was apocalyptic in nature (Lynch, 1999). Further, a definite date could be pinpointed as the danger point - January 1, 2000 – a date which provided a "time bomb" for a number of apocalyptic predictions (Yourdon & Yourdon, 1998).

Some religious groups, though, saw the Y2K problem as a symptom of a broader, religious-based apocalyptic event, with some proclaiming that the

---

Author info: Correspondence should be sent to: Larry Powell, Communication Studies Department, University of Alabama at Birmingham, 901 South 15th Street, Birmingham, AL 35294-2060.

North American Journal of Psychology, 2001, Vol. 3, No. 2, 295-302.

© NAJP

Y2K bug would signal the end of the world (Larue, 1999). As one religious adherent said, "The fingerprints of God are on this coming storm" (Farrar, 1999, p. 14). The approaching date became a frequent topic for Sunday sermons (Feldstein, 1999), leading the American Banking Association to provide a "model" Y2K sermon to churches that offered reassurance ("Bankers turn to religion...., 1999).

Such attitudes should not be surprising, given the important role that religion can have in attitude formation. Religious values are a particularly potent basis of value expressive attitudes because religious beliefs are often related to both self-perception and group identification (with a church or religious organization). Group identification is a known factor in framing perceptions. Penning (1988) and Katz (1960, p. 174) argued that identification with a religious organization could influence an individual to internalize the values of the organization. That concept was supported by Wald (1987) who found that integration into religious groups influenced both the political attitudes and political behavior of those groups' members.

Religious values fall within Katz's (1960) value-expression function of attitudes in which "the individual derives satisfactions from expressing attitudes appropriate to his personal values and to his concept of himself" (p. 170). In essence, value-expressive attitudes reflect individuals' self-perception of who they are and their role within an external world. They allow the individual to give "...positive expression to his central values and to the type of person he conceives himself to be" (p. 173).

Research has frequently supported the contention that religion can influence secular attitudes on such topics as divorce, abortion, school curricula, feminism, and sexual liberation (Hammond, Shibley & Solow, 1994). Jalen (1998) demonstrated that religious affiliations and beliefs affect a wide variety of political attitudes and activities. Similarly, Leege and Kellstedt (1993) found that religion influenced the political behavior of large groups of voters, and Lopatto (1985) argued that theology based on religious affiliation served as a sieve through which voters filtered and evaluated their political world. That phenomenon was a major factor during the 1980s with the participation of the Moral Majority (Wald, 1987) and continued during the 1990s (Manza & Brooks, 1997) through groups like the Christian Coalition.

The Y2K crisis had the potential to trigger responses based on religious values. The apocalyptic nature of the predicted problem was tailor-made for apocalyptic religions, while the machine-based source of the problem could trigger questions about human reliance on the "god" of technology. Such possibilities create a number of potential responses on the part of active church-goers that could (1) affect their knowledge of Y2K, (2) affect their perception of the seriousness of the problem, (3) make them more sensitive to information in the media about Y2K, (4) generate more interpersonal discussions about Y2K (Swanson, 1976), and (5) lead them to make more preparations in readiness for Y2K. Thus five research hypotheses were investigated.

H_1: Frequent church-goers will be more likely than non-church goers to know the meaning of the term "Y2K."

H_2: Frequent church-goers will perceive the Y2K crisis to be more serious than non-church goers.

H_3: Frequent church-goers will report seeing more information about Y2K in the media than will non-churchgoers.

H_4: Frequent church-goers will have more interpersonal discussion about Y2K than will non-churchgoers.

H_5: Frequent church-goers will make more preparations for Y2K than will non-churchgoers.

METHOD

Participants

The participants were 497 residents of Alabama (209 men, 288 women; median age, 49.2), who were interviewed in a stratified random sample in November of 1999 (margin of error, ± 4. percent). The racial make-up of the sample (82.3% white, 16.1% African-American, 1.6% other) was representative of the state's racial diversity.

Instrument

Hypothesis one was operationalized with the following question: "Some people have mentioned a potential problem when the year changes, something they call the 'Y-2-K bug.' Off hand, would you happen to know what "Y-2-K" stands for?" If the participant identified Y-2-K as "Year 2000," they were given credit for a positive response to this question. Other responses were coded as incorrect.

Perceived seriousness of the problem was measured by asking the following question: "Overall, would you say that the potential for computers to crash and cause major problems when the year 2000 arrives is (1) a very serious problem, (2) something to be concerned about, but not a major problem, or (3) not much of a problem at all? (4) unsure." Those who chose "unsure" were excluded from the analysis.

Perception of information in the media was measured with the following question: "How much have you heard about the Y-2-K problem in the media, (1) a lot, (2) some, (3) not very much, or (4) not at all."

Level of interpersonal discussions about Y2K was measured with the following question: "How much have you discussed the Y-2-K problem with your friends and neighbors, (1) a lot, (2) some, (3) not very much, or (4) not at all?"

Preparation for the crisis was measured with the following question: "The biggest concern for most people about the changing of the new year is that computers may shut down, causing problems for banks, power companies and transportation. Have you or your family made any plans to handle that in case it happens? (1) yes (2) no (3) unsure." Those who chose "unsure" were excluded from the analysis.

Church attendance was measured by asking" Do you usually attend church once a week or more, a few times a month, or a few times a year or less?" Those who said they attended church once a week or more were defined as "frequent churchgoers." Those who said they attended "a few times a month" were defined as "occasional churchgoers." Those who said they attended church a few times a year or less were defined as "non-churchgoers."

Procedure

Participants were selected using the following procedure. First, a sampling formula was developed by dividing the state into six geographical regions, and percentages were assigned to each area based upon voter participation in the 1996 presidential election. Second, 8,000 names matched with telephone numbers were randomly selected in clusters of twenty names, and these were apportioned into the six geographical areas. Third, names were randomly selected from that list, contacted by telephone, and asked if they would participate in the survey. If they refused, a second number was called (alternately the one immediately above or below the original number selected). Participants were monitored to maintain gender balance in each region. A total of 586 respondents were contacted by phone, of which 497 agreed to participate in the survey (response rate, 85%).

Hypothesis 1 was tested with a 3 x 2 (church attendance by knowledge) chi-square analysis. Hypothesis 2 was tested with a 3 x 3 (church attendance by seriousness rating) chi-square analysis. Hypothesis 3 was tested with a 3 x 4 (church attendance by media information) chi-square analysis. Hypothesis 4 was tested with a 3 x 4 (church attendance by interpersonal discussion) chi-square analysis. Hypothesis 5 was tested with a 3 x 2 (church attendance by preparations) chi-square analysis.

RESULTS

Hypothesis 1 was not accepted. There was no significant relationship ($X^2_2 = 3.41$) between knowing the meaning of Y2K and church attendance.

Hypothesis 2 was not accepted. Perceived seriousness of the Y2K issue did not vary ($X^2_4 = 8.51$) by church attendance.

Hypothesis 3 was not accepted, but an unanticipated result was identified. Church attendance was associated with perceptions of the amount of information about Y2K in the media ($X^2{}_6 = 29.37$, $p < .001$), but not in the predicted manner. Occasional church-goers were more likely to report that the media had "a lot" of information about Y2K (77%) than did frequent (62%) and non-church-goers (67%).

Hypothesis 4 was accepted. Frequency of interpersonal discussions varied by church attendance ($X^2{}_6 = 21.19$, $p < .001$). Frequent church-goers had more interpersonal discussions about Y2K (50%) than did non-church-goers (35%).

Hypothesis 5 was accepted. Those who had taken preparations against a Y2K problem varied in terms of church attendance ($X^2{}_2 = 12.59$, $p < .001$), with frequent church-goers making more preparations (30%) than non-church-goers (24%).

DISCUSSION

First, it should be noted that one of our significant differences (a 6 % difference in hypothesis 5) is rather small and is partly due to large sample size. The others are not especially large, so we recommend a cautious interpretation of the results.

Religious activity was associated with perceptions of the amount of media coverage of Y2K, frequency of interpersonal discussion on the topic, and level of preparations for the possible crisis. Frequent church-goers were more likely to have taken preparations against a Y2K problem and engaged in more interpersonal discussions on the topic than did non-churchgoers, while occasional church-goers reported seeing more information about Y2K in the media.

No significant relationships were identified for either knowledge of the term "Y2K" or perception of the salience of the issues. Thus, although church-goers saw more information in the media and were more likely to have discussed the Y2K problem with others, they were no more knowledgeable about it than were there non-church-going counterparts, nor did they see the problem as being more serious than non-church-goers.

The lack of variation in the knowledge factor may be due to the simplistic test used in this survey. All that was required to pass this test was an awareness that Y2K was an abbreviation for "Year 2000," and that level of knowledge could have been achieved with a minimal level of exposure to the information. Indeed, 60% of our participants knew the meaning of the term "Y2K." Thus, this finding may be an artifact of the measuring technique, since it does not provide a means of interpreting multiple levels of knowledge.

Of more interest is the finding that church-goers were more likely to have taken preparations for Y2K, even though they did not see the problem as being any more serious than did non-church-goers. Two factors may contribute to this result, one that is topic-specific in nature and one that might be a generalized trait for those who are active church-goers. First, on a topic-specific level, it may be a by-product of the perceived link between the apocalyptic nature of the perceived crisis and the apocalyptic theology of some Christian religions. Such a perceived link could have increased the motivation for active church-goers to engage in behavior related to their attitudes on this topic. Second, on a generalizable trait level, it may indicate that church-goers are more likely to act on the basis of their attitudes than non-church-goers. One theological element in many American religious organizations is the belief that religious values should be manifested in behaviors (Weber, 1958). If an individual has such a religious mentality, the tendency to act on the basis of one's beliefs may extend to non-religious attitudes.

Regardless, the results indicate that religious activity is associated with behavioral and perceptual elements related to the perceived Y2K crisis. Future studies should be conducted to extend knowledge on the relationship between attitudes, perceptions, and behaviors of those who are active in the religious community.

REFERENCES

Bankers turn to religion in effort to ally Y2K fears (1999, Aug. 26). *St. Louis Post-Dispatch*, p. A1.

Farrar, S. (1999). *Spiritual survival during the Y2K crisis.* Nashville: Thomas Nelson.

Feldstein, F. (1999). Multimedia pedagogy and Sunday morning millennial fever. *College Literature, 26(3)*, 205-239.

Gantz, W. (1983). The diffusion of news about the attempted Reagan assassination. *Journal of Communication, 33*, 56-66.

Greenberg, B.S. (1964). Diffusion of news of the Kennedy assassination. *Public Opinion Quarterly, 28*, 225-232.

Hammond, P.R., Shibley, Mark A., & Solow, P.M. (1994). Religion and family values in presidential voting. *Sociology of Religion, 55*, 277-290.

Hickson, M., Powell, L., & Bodon. J. (2001). Rejection of crisis information: Public apathy and the macro-crisis of the Y2K. Paper for the Southern States Communication Association. Lexington, KY.

Jalen, T.G. (1998). Research in religion and mass political behavior in the United States: Looking both ways after two decades of scholarship. *American Politics Quarterly, 26*, 110-134.

Katz, D. (1960). The functional approach to the study of attitudes. *Public Opinion Quarterly, 24*, 163-204.

Kraus, S., Davis, D., Lang, G.E., & Lang, K. (1976). Critical events analysis. In S.H. Chaffee (Ed.), *Political communication: Issues and strategies for research* (pp. 195-216). Beverly Hills, CA: Sage.

Larue, G.A. (1999). The Bible and the prophets of doom. *Skeptical Inquirer, 23(1)*, 25-32.

Leege, D.C., & Kellstedt, L.A. (1993). *Rediscovering the religious factor in American politics.* Armonk, NY: M.E. Sharpe.

Lopatto, P. (1985). *Religion and the presidential election.* New York: Praeger.

Lynch, A. (1999). The millennium thought contagion: Is your mental software Year 2000 compliant? *Skeptical Inquirer, 23(6)*, 32-37.

Manza, J., & Brooks, C. (1997). The religious factor in U.S. presidential elections, 1960-1992. *American Journal of Sociology, 103*, 38-81.

Mendelsohn, H. (1964). Broadcast vs. personal sources of information in emergent public crises: The presidential assassination. *Journal of Broadcasting, 8*, 147-156.

Penning, J.M. (1988). The political behavior of American Catholics: An assessment of the impact of group integration vs. group identification. *Western Political Quarterly, 41 (2)*, 289-308.

Riffe, D., & Stovall, J.G. (1989). Diffusion of news of shuttle disaster: What role for emotional response? *Journalism Quarterly, 66*, 551-560.

Steinfatt, T.M., Gantz, W., Siebold, D.R., & Miller, L.D. (1973). News diffusion of the Wallace shooting: The apparent lack of interpersonal communication as an artifact of delayed measurement. *Quarterly Journal of Speech, 59*, 401-412.

Swanson, D.L. (1976). Information utility: An alternative perspective in political communication. *Central States Speech Journal, 27*, 95-101.

Wald, K.D. (1987). *Religion and politics in the United States.* New York: St. Martin's Press

Weaver-Lariscy, R.A., Sweeney, B., & Steinfatt, T. (1984). Communication during assassination attempts: Diffusion of information in attacks on President Reagan and the Pope. *Southern Speech Communication Journal, 49,* 258-276.

Weber, M. (1958). *The Protestant ethic and the spirit of capitalism.* New York: Charles Scribner.

Yourdon, E., & Yourdon, J. (1998). *What the Year 2000 computer crisis means to you: Time bomb 2000.* Upper Saddle River, NJ: Prentice-Hall.

Stability of Self-esteem as Moderator of the Relationship Between Level of Self-esteem and Depression

A. F. de Man
B. I. Becerril Gutiérrez
N. Sterk
Bishop's University

One-hundred-and-thirty-one undergraduate university students participated in a study of the relationship between level and (in)stability of self-esteem and depression. Results of correlational and stepwise multiple regression analyses showed that depression was significantly related to low as well as to unstable self-esteem. Furthermore, for individuals with high self-esteem, variation in self-esteem stability did not have a significant moderating influence, whereas for those with low self-esteem, stable self-esteem appeared to be a protective factor.

Within diverse theoretical orientations, self-esteem plays an important role in the etiology and maintenance of depression. Abraham (1911) and Freud (1917) suggested that anger turned inward upon the self generates low self-esteem and depression. Fenichel (1945) described depressives as individuals who try to compensate for their depleted self-worth by seeking comfort and reassurance from others. Bibring (1953) claimed that depression results when the ego feels helpless at achieving its goals and experiences a loss of self-esteem. Beck (1967) suggested that negative self-evaluations as part of a negative cognitive triad play an important role in depressive episodes. Seligman (1975) and Abrahamson, Metalsky, and Alloy's (1989) respectively indicated that learned helplessness and hopelessness may result in a decline in self-esteem, and depression. Hence, there seems to be agreement that self-esteem is an important correlate of depression, which is reflected in the fact that the DSM IV (American Psychiatric Association, 1994) lists negative self-esteem as one of the diagnostic criteria for depression.

Many studies have shown a strong negative relationship between self-esteem and depression (e.g., Brockner & Guare, 1983; Brown, Bifulco, Harris, & Bridge, 1986; Hammen, Marks, deMayo, & Mayol, 1985; Tennen & Herzberger, 1987). However, most of these investigations have

Author info: Correspondence should be sent to: A. F. de Man, Department of Psychology, Bishop's University, Lennoxville, Quebec, Canada J1M 1Z7 (e-mail: ademan@ubishops.ca).

North American Journal of Psychology, 2001, Vol. 3, No. 2, 303-308.

© NAJP

not taken into consideration Rosenberg's (1986) suggestion that one may distinguish between *level* of self-esteem (i.e., the amount of enduring global sense of worth and well-being a person has) and *stability* of self-esteem (i.e., the magnitude of short-term fluctuations in self-esteem level). Most have focused almost exclusively on level of self-esteem as the element that is associated with depression, and have disregarded the possible effect of variation in stability.

Kernis, Grannemann and Mathis (1991) investigated the relationship between level of self-esteem and depression in individuals with stable and unstable self-esteem. Consistent with previous research, they found an inverse relationship between overall level of self-esteem and depression. Further analysis showed that instability was related to higher levels of depression for individuals with high self-esteem and to lower levels of depression for those with low self-esteem. Roberts and Monroe (1992), on the other hand, found evidence for labile self-esteem as a vulnerability factor for depressive symptoms, with unstable self-esteem predicting increases in depressive symptoms in high as well as in low self-esteem individuals.

In light of the discrepancy between the findings of Kernis *et al.* (1991) and Roberts and Monroe (1992), the present study aimed to clarify the relationship between level of self-esteem, (in)stability of self-esteem, and depression.

METHOD

Participants

Participants were an availability sample of 131 undergraduate university students (estimated age range: 18 - 24). This non-clinical group included 108 women and 23 men; five were non-Caucasian.

Instruments and Procedures

Self-esteem level and self-esteem (in)stability were assessed by Rosenberg's (1965) Self-Esteem Scale. This scale, a well-validated measure of global self-regard (Demo, 1985; Wylie, 1974), consists of 10 items rated on 4-point scales ranging from strongly agree to strongly disagree. Scores may vary from 10 (low self-esteem) to 40 (high self-esteem).

Respondents completed the Self-Esteem Scale during regular class periods on five different occasions (approximately equal intervals) over a five-week period. Each time, they were instructed to circle the response options that best reflected how they felt at that particular moment. Level of self-esteem of the individual participants was obtained by calculating the mean of the individual's total scores across the five assessments. Self-esteem (in)stability was computed as the standard deviation of the respondent's total scores across the repeated measurements (Kernis *et al.*, 1991), with higher scores reflecting greater instability.

Depression was measured by the Beck Depression Inventory (Beck, Ward, Mendelson, Mock, & Erbauch, 1961) which has been shown to be a valid measure of depressive symptoms among college students (Bumberry, Oliver, & McClure, 1978). The 21 items of this scale represent different manifestations of depressive states. Each item consists of 4 or 5 statements reflecting various levels of the depressed feeling in question. Numerical values for each item may range from 0 (neutral) to 3 (maximum intensity); total scores may range from 0 (low depression) to 63 (high depression). The depression scale was administered in the fifth week following completion of the last self-esteem questionnaire.

RESULTS AND DISCUSSION

Level of self-esteem scores varied from 20.20 to 40 ($M = 30.87$; $SD = 4.45$), and (in)stability scores (i.e., standard deviations) ranged from 0 to 7.19 ($M = 2.00$; $SD = 1.20$). Depression scores varied from 0 to 30 ($M = 7.49$; $SD = 6.18$).

Skewed distributions were found for some of the variables. Transformations were performed and bivariate correlations for the original as well as the transformed data were calculated. The transformed data behaved no better than the original ones. Therefore, subsequent analyses were limited to the original scores. Furthermore, correlation analyses showed no significant relations between gender and the variables under consideration. Accordingly, gender was omitted from subsequent analyses.

Kernis *et al.* (1991) and Roberts and Monroe (1992) reported a significant negative relationship between level and (in)stability of self-esteem (with level of self-esteem respectively measured as self-esteem average or aggregated self-esteem), suggesting greater instability at the low self-esteem level. In the present study, these two elements of self-esteem were not significantly related ($r = -.12$, $p < .17$), indicating that varying levels of self-esteem were randomly associated with varying levels of stability and instability.

Correlation analyses further showed that low self-esteem ($r = -.54$, $p < .0001$) and unstable ($r = .25$, $p < .003$) self-esteem each were significantly related to depression. Similar findings were reported by Kernis *et al.* (1991) and Roberts and Monroe (1992).

A stepwise multiple regression analysis with depression as the dependent variable and level of self-esteem, self-esteem (in)stability, and the interaction between the latter two variables as predictors, identified level of self-esteem as best predictor, accounting for 29.4% ($F_{1,129} = 53.74$, $p < .0001$) of the variance in depression. This percentage could be significantly increased to 33.0% ($F_{2,128} = 31.55$, $p < .0001$) and 35.4% ($F_{3,127} = 23.19$, $p < .0001$) respectively, by including self-esteem (in)-stability, and Level by (in)stability interaction in the predictive equation.

Consistent with prior research, the significant main effect of level of self-esteem indicated that respondents with low self-esteem generally reported greater depressive tendencies. The significant main effect of self-esteem (in)stability showed that participants with unstable self-esteem tended to report similar depressive inclinations. More important, the interaction between self-esteem level and self-esteem (in)stability was significant. To interpret this interaction, mean depression scores were calculated for four groups of participants created by taking the upper 30% and lower 30% of the respective scores for level and (in)stability of self-esteem (see Table 1).

TABLE 1 Mean Scores and Standard Deviations for Depression in Relation to Level and Stability of Self-esteem

	Stability of Self-esteem			
	Unstable		Stable	
	M	SD	M	SD
Low Self-esteem Level	15.88	6.69	9.33	6.00
High Self-esteem Level	4.87	5.48	3.41	3.10

A survey of these mean scores showed that the group of respondents with low, unstable self-esteem had the highest mean depression score, whereas the lowest mean depression score was found among participants with stable, high self-esteem. A distribution such as this would not be congruent with the findings reported by Kernis *et al.* (1991), but would be consistent with the observations made by Roberts and Monroe (1992).

T-tests revealed a significant difference between the stable and unstable *low* self-esteem groups ($t = 2.86$, $p < .008$), but not between the stable and unstable *high* self-esteem ones ($t = .94$, $p < .35$). Further differences were found between the *stable* low and high self-esteem groups ($t = 3.57$, $p < .001$), and between the *unstable* low and high self-esteem ones ($t = 4.99$, $p < .0001$), with the difference between the two unstable groups being significantly greater than the difference between the two stable ones.

These findings present a virtual mirror image of the results reported by Kernis *et al.* (1991). The latter found a substantial difference in depression between low and high self-esteem scorers, but only for individuals with *stable* self-esteem, whereas the present study found significant differences in depression between the two levels of self-esteem both for respondents with *stable* self-esteem and those with *unstable* self-esteem, but with a significantly larger difference among individuals with *unstable* self-esteem. Furthermore, Kernis *et al.* reported a substantial difference in depression between respondents with stable and unstable *high* self-esteem, and a minimal difference between individuals with stable and unstable *low* self-esteem, whereas in the present study a significant difference in depression was found between participants with stable and unstable *low* self-esteem, but not between respondents with stable and unstable *high* self-esteem. Hence, the findings of the present study are not congruent with those reported by Kernis *et al.*

The study by Roberts and Monroe (1992) demonstrated the viability of unstable self-esteem increasing depressive symptoms. The findings of the present study seem to agree with this. However, the cumulative interaction between level and (in)stability of self-esteem is only apparent at the *low* self-esteem level. For individuals with high self-esteem, variation in self-esteem stability does not appear to have a significant moderating influence, whereas for those with low self-esteem, stability may serve a protective function. Stated in other words, for individuals with low self-esteem, instability of self-esteem should be a good predictor of greater depression.

REFERENCES

Abraham, K. (1911). Notes on psychoanalytic investigation and treatment of manic-depressive insanity and applied conditions. In *Selected papers of Karl Abraham, M.D.* (D. Bryan & A. Strachey, Trans.). London: Hogarth Press (1948).

Abrahamson, L., Metalsky, G., & Alloy, L. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review, 96*, 358-372.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects.* New York: Harper & Row.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbauch, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*, 53-56.

Bibring, E. (1953). The mechanism of depression. In P. Greenacre (Ed.), *Affective disorders*. New York: International Universities Press.

Brockner, J., & Guare, J. (1983). Improving the performance of low self-esteem individuals: An attributional approach. *Academy of Management Journal, 26*, 642-656.

Brown, G. W., Bifulco, A., Harris, T., & Bridge, L. (1986). Life stress, chronic subclinical symptoms and vulnerability to clinical depression. *Journal of Affective Disorders, 11*, 1-19.

Bumberry, W., Oliver, J. M., & McClure, J. N. (1978). Validation of the Beck Depression Inventory in a university population using psychiatric estimate as the criterion. *Journal of Consulting and Clinical Psychology, 46*, 150-155.

Demo, D. H. (1985). The measurement of self-esteem: Refining our methods. *Journal of Personality and Social Psychology, 48*, 1490-1502.

Fenichel, O. (1945). *The psychoanalytic theory of neurosis*. New York: Norton.

Freud, S. (1917). Mourning and melancholia. In J. Strachey (Ed. and Trans.), *The complete psychological works* (Vol. 16).New York: Norton (1976).

Hammen, C., Marks, T., deMayo, R., & Mayol, A. (1985). Self-schemas and risk for depression: A prospective study. *Journal of Personality and Social Psychology*, 49, 1147-1159.

Kernis, M. H., Grannemann, B. D., & Mathis, L. C. (1991). Stability of self-esteem as a moderator of the relation between level of self-esteem and depression. *Journal of Personality and Social Psychology*, 61, 80-84.

Roberts, J. E., & Monroe, S. M. (1992). Vulnerable self-esteem and depressive symptoms: Prospective findings comparing three alternative conceptualizations. *Journal of Personality and Social Psychology*, 62, 804-812.

Rosenberg, M. (1986). Self concept from middle childhood through adolescence. In J. Suls & A. Greenwald (Eds.), *Psychological perspectives on the self* (Vol. 3). Hillsdale, NJ: Erlbaum.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton Univ. Press.

Seligman, M. E. P. (1975). *Helplessness: On depression, development, and death*. San Francisco: Freeman.

Tennen, H., & Herzberger, S. (1987). Depression, self-esteem, and the absence of self-protective attributional biases. *Journal of Personality and Social Psychology*, 52, 72-80.

Wylie, R. C. (1974). *The Self-concept* (Vol. 1). Lincoln, NE: University of Nebraska Press.

Legal Attitudes Scale: A Factor Analytic Study

James V. Couch and Donna L. Sundre
James Madison University

A factor analysis of responses to the Legal Attitudes Scale identified three independent factors. Preliminary construct and content validity results are reported for these three factors. Results are discussed in terms of the continued development of a scale to measure attitudes in prospective jurors.

The purpose of this paper is to demonstrate the important contribution psychology can provide in assessing jury beliefs. Singer (1995) provided jury selection guidance for practicing attorneys when he stated that a juror's beliefs are "the only consistent predictors of jury verdict" (p. 18). Despite this sage advice, the literature is replete with conflicting advice on how best to undertake the important task of jury selection.

When a person is a criminal defendant or two parties are involved in a civil suit, a trial by an impartial jury is a right guaranteed by the U. S. Constitution. In exercising this right, counsel for both sides try to select a jury that meets the condition of impartiality. A variety of sources and diverse methods have been identified and accessed to aid counsel in undertaking this important task. Numerous trial advocacy texts give guidance on jury selection (Jeans, 1993; Vinson, 1993). Trial folklore suggests that individuals of certain occupations are more favorable to the defense or prosecution. Fahringer (1993) counsels attorneys that attention to a prospective juror's way of dress, facial expressions, body language, and the way the juror uses their eyes are important clues that will help the attorney "achieve our objectives efficiently and effectively" (p. 204). Presumably this means selecting a jury which is predisposed to the attorney's position. Fulero & Penrod (1990) provide an entertaining review of myths that have abounded within the legal community as to demographic variables that have been suggested as appropriate variables to use in selecting jurors. As these authors point out there are many conflicting suggestions arising from these myths. Schutte (1996) comes to a similar conclusion, namely that the use of demographic variables in the selection of jurors is fraught with error.

Psychologists can provide contributions based upon science to assist legal counsel in fulfilling these critical duties. First, and foremost, the field of psychology has defined and developed methods by which stable

---

Author info: Correspondence should be sent to: Donna L. Sundre, Center for Assessment and Research Studies, MSC 6806, James Madison University, Harrisonburg, Virginia 22807; e-mail: sundredl@jmu.edu.

North American Journal of Psychology: 2001, Vol. 3, No. 2, 309-320.

© NAJP

personality traits can be assessed. Many of these assessment methods have been scrutinized psychometrically and researched over time to provide reliable and valid measures of personality constructs thought to be associated with consistent juror beliefs and verdicts.

Personality factors have been the focus of attention as counselors have tried to select a jury that will be impartial but favorably predisposed to their side of the litigated issue. One of the personality factors which has been investigated is that of the authoritarian personality. Narby, Cutler, & Moran (1993) reported the results of a meta-analysis that suggests that authoritarianism is a reliable predictor of juror behavior. Boehm (1968) developed an instrument called the Legal Attitudes Questionnaire (LAQ), which purports to measure authoritarian, anti-authoritarian, and equalitarian tendencies. The original LAQ consisted of 30 items that were arranged in triads of one authoritarian, one anti-authoritarian, and one equalitarian item. Boehm defined the three LAQ personality tendencies as follows:

> Authoritarian items expressed right wing philosophy, endorsed indiscriminately the acts of constituted authority, or were essentially punitive in nature. Anti-authoritarian items expressed left wing sentiments, implied that the blame for all antisocial acts rested with the structure of society, or indiscriminately rejected the acts of constituted authority. Equalitarian items endorsed traditional, liberal, nonextreme positions on legal questions or were couched in a form that indicated the questions reasonably could have two answers (p. 740).

To evaluate the LAQ in a mock jury format, Boehm (1968) presented participants with one of two cases. One case was written such that the "correct" verdict was acquittal (Not Guilty scenario) while the second case called for a verdict of either manslaughter or second-degree murder (Guilty scenario). Participants were classified into age groups with "younger" participants being those less than 21 years old and "older" ones being those 21 years of age or older. Results indicated that both age group participants judged to be high on the authoritarianism dimension were more likely to convict a person described in the Not Guilty scenario. Older participants judged to be high on the anti-authoritarianism dimension were more likely to vote for acquittal of the person described in the Guilty scenario. The results for younger anti-authoritarian participants given the Not Guilty scenario were inconclusive. It should be noted that Boehm's classification of items on the LAQ was arbitrary and not governed by any statistical criterion such as factor analysis.

Other studies (Cowan, Thompson, & Ellsworth, 1984; Moran & Comfort, 1982) have demonstrated results similar to Boehm (1968), with participants scoring high on the authoritarian dimension being likely to return a verdict of guilty. In the Moran & Comfort study it should be noted that only 24 of the original 30 LAQ items were employed, and requesting participants to rate levels of agreement with the items replaced the forced-choice format. They also scored each item as legal authoritarian versus antiauthoritarian.

Kravitz, Cutler, & Brock (1993) examined the psychometric properties of Boehm's LAQ. Kravitz, et al. believed that the forced choice nature of the items of the original LAQ led to a number of psychometric problems. To this end, they changed the forced choice option to a 6-point Likert-format response scale. In addition, Kravitz, et al. reduced the original 30-item scale to a 23-item scale (RLAQ23). They provided some construct validity evidence through moderately high correlations with measures of death penalty attitudes and a revised version of the F-scale authoritarianism scale. A low correlation with a social desirability scale was also reported. In two different studies, one with college students as participants and the second with jury-eligible individuals, Kravitz, et al. demonstrated the superiority of the RLAQ23 as well as the usefulness of using the Likert-format response scale.

Kravitz, et al. reported performing a "confirmatory factor analysis … (which) revealed that RLAQ23 was not unidimensional" (p.672). Further they reported that "… exploratory factor analyses varying in methods of factor extraction and rotation failed to reveal factors that were clearly interpretable and reliable" (p. 672). Even though Kravitz, et al. reported performing these factor analytic investigations of the RLAQ23, they failed to report the results of such investigations and relied upon Boehm's original classification of items when developing scores for the authoritarian, anti-authoritarian, and equalitarian scales.

The present investigation seeks to extend the work of Kravitz, et al. by reporting a rotated factor structure for the RLAQ23. Factor scores for the resulting factors can be obtained by averaging the subject's response to each item comprising the factor (as was done in Kravitz, et al.), by totaling responses to each item comprising the factor, or by using the factor loadings to develop a weighted factor score. Further, by examining the factor structure of the RLAQ23 it is possible to determine whether this scale measures just the three dimensions as suggested by Boehm (1968) or whether the RLAQ23 items are measuring other aspects of the subject's underlying belief system.

METHOD

Participants

A total of 225 students from the University subject pool volunteered to participate in the study for course credit. Forty-four male and 181 female participants completed the survey instruments. The vast majority of the students (88%) were eighteen or nineteen years of age. A variety of majors were represented, specifically Arts and Letters (24%), Education and Psychology (33%), Business (23%), Science and Mathematics (11%), and Integrated Science and Technology (9%).

TABLE 1 Demographic Questions and Response Options [a]

Question	Response Options
How would you classify your political orientation in terms of the two major political parties?	1 – Strong Democrat to 5 – Strong Republican
How do you think of yourself?	1 – Extremely liberal to 5 – Extremely conservative
Do you favor or oppose the death penalty for murder?	1 – Favor 2 - Oppose
Do you favor or oppose the requirement that parents have to be notified before an abortion can be performed on their minor daughter?	1 – Favor 2 - Oppose

[a] Several other demographic questions were asked but were not a part of the current data analysis.

Materials

Upon arriving for the study, participants were read a set of standardized instructions. After hearing the instructions all of them were given an informed consent document, which they signed and returned to the experimenters.

Each subject received a packet of instruments containing a demographic sheet, a copy of the Legal Attitude Questionnaire (RLAQ23), and an Occupational Survey. Administration of the Occupational Survey was for another investigation and those results will not be reported as part of the present investigation.

The demographic sheet contained questions concerning the participant's age, gender, and college in which they were majoring. In addition, four questions were included to be used in construct validation of the RLAQ23 (see Table 1).

The RLAQ23 consisted of the 23-item revision of Boehm's (1968) LAQ that was recommended for use by Kravitz, et al. (1993). Participants used a 6-point scale ranging from Strongly Disagree (1) to Strongly Agree (6). Fourteen of the RLAQ23 items were reversed scored as directed by Kravitz, et al. prior to forming scores.

TABLE 2 Factor Structure for the RLAQ23 [a]

Question	Factor 1 Loading (Libertarian)	Factor 2 Loading (Authoritarian)	Factor 3 Loading (Police)
Unfair treatment of underprivileged groups and classes is the chief cause of crime.	.61		
Too often, minority group members don't get fair trials.	.60		
In the long run, liberty is more important than order.	.59		
In determining a person's guilt or innocence, existence of a prior arrest record should not be considered.	.52		
When there is a "hung" jury in a criminal case, the defendant should always be freed and the indictment dismissed.	.47		
No one should be convicted of a crime on the basis of circumstantial evidence, no matter how strong such evidence.	.42		
Wiretapping by anyone and for any reason should be completely illegal.	.41		
Search warrants should clearly specify the person or things to be seized.	.36		
A society with true freedom and equality for *all* would have very little crime.	.36		
There is no need in a criminal case for the accused to prove his/her innocence beyond a reasonable doubt.	.35		
It is moral and ethical for a lawyer to represent a defendant in a criminal case even when he/she believes the client is guilty.	.31		
Defendants in a criminal case should be required to take the witness stand.		.73	
The accused should be made to take lie detector tests.		.62	
Evidence illegally obtained should be admissible in court if such evidence is the only way of obtaining a conviction.		.52	
The law coddles criminals to the detriment of society.		.51	
Too many obviously guilty persons escape punishment because of legal technicalities.		.49	
Citizens need to be protected against excess power as well as against criminals.			.63
Police should be allowed to arrest and question suspicious looking persons.			.60
The freedom of society is endangered as much by overzealous law enforcement as by the acts of individual criminals.			.52
Upstanding citizens have nothing to fear from police.			.47

[a] Three questions did not meet the .30 loading criterion. The question "It is better for society that several guilty persons be freed than one innocent person wrongly imprisoned" had a factor loading or .27 on Factor 1. The question "Any person who resists arrest commits a crime." Had a factor loading of .18 on Factor 2. The question "Because of the oppression and persecution many minority group members suffer, they deserve leniency and special treatment in the courts." Had a factor loading of -..29 on Factor 3. Some questions have been abbreviated slightly.

RESULTS AND DISCUSSION

All statistical analyses were performed using the entire data set with exclusion of missing data. If data were missing on any item of the RLAQ23, the data for that subject was excluded from analysis on that instrument.

Demographic Variables

Chi-square tests of independence were performed using various combinations of the demographic variables. Prior to conducting the tests, the response alternatives for the Liberal/Conservative Orientation demographic question were collapsed into three categories by grouping the alternatives, Extremely Liberal and Liberal, into one category, the alternatives Extremely Conservative and Conservative into another category, and retaining the original Moderate response category. In like fashion, the response alternatives for the Political Orientation demographic question were collapsed into three categories labeled Democratic Orientation, Independent, and Republican Orientation. The categories were collapsed because the extreme categories had too few observations.

Of particular interest to the present investigation is the information from the political orientation and liberal/conservative orientation questions, since these items were used to test for construct validity of the RLAQ23. Participants who identified themselves with the Democratic Party were more likely to identify themselves as liberal while those who identified with the Republican Party were more likely to identify themselves as conservative, $\chi^2(_4, n = 223) = 67.87, p < .001$. Participants who identified with the Republican Party were more likely to favor the death penalty for murder than participants who identified with the Democratic Party, $\chi^2(_2, n = 223) = 10.49, p < .005$. Most participants tended to favor the notification of parents in the case of a minor seeking an abortion however fewer "conservatives" than would be expected by chance opposed such notification, $\chi^2(_2, n = 224) = 9.02, p = .011$. In summary, these findings parallel those of Kravitz, et al. (1993), in which participants reporting conservative political identification were more likely to favor the death penalty. These relationships were used in the construct validation of the RLAQ23.

Psychometric Characteristics of the RLAQ23

Across the total set of 23 items the potential range for the total scale was 23 to 138 with higher scores indicating greater authoritarian or right wing views. The mean score was 85.63 with a standard deviation of 10.31. The internal consistency coefficient, estimated using Cronbach's alpha, was .63. This is similar to the reliability reported by Kravitz, et al.

(1993) in their study using undergraduate college students. It is interesting to note that when the scale was administered to a sample of jury-eligible adults the reliability increased dramatically, even though the sample size consisted of only about 100 participants. We believe this reliability increase is due to the increase in variability within the jury-eligible sample. The Kravitz, et al. (1993) paper did not report the descriptive statistics for the total scale.

The hypothesized unidimensional structure of the RLAQ23 was tested using Amos Version 3.51 (Arbuckle, 1995). This structural equation modeling (SEM) confirmatory factor analysis (CFA) result paralleled the result earlier obtained by Kravitz, et al. (1993). The unidimensional structure was not well supported by the current study sample data, χ^2 (230, n = 211) = 483.72, $p < .001$. In other words, the discrepancy between the hypothesized structure and the observed covariance matrix was quite large. However, the chi-square statistic has been shown to be quite sensitive to sample size (Loehlin, 1992), such that relatively meager discrepancies lead to rejection with large samples and major discrepancies test non-significant with small samples. Two other fit indices are presented below that also refute the unidimensional model. It should be noted that the Tucker-Lewis (1973) index has been shown through real and simulated data to be relatively independent of sample size in confirmatory factor analysis studies (Marsh, Balla, & McDonald, 1988). Both the Tucker-Lewis Index (TLI) = .370 and the Bentler's (1990) Comparative Fit Index (CFI)= .427 provide strong evidence that the sample data do not support the hypothesized single factor structure. These indices follow a 0 to 1 scale with values greater than .9 considered indicators of good fit. These indices further corroborate the earlier results suggesting lack of fit. At this stage of analysis, an exploratory factor analysis (EFA) was conducted. While SEM is considered *a priori*, methods termed 'exploratory' are often employed to test *a priori* hypotheses. Alternative *a priori* models are often identified and tested. For the current study, the unidimensional model was specified, then tested in comparison to an alternative dimensional model. Alternative model generation can be specified but only when theory and statistical correspondence provide appropriate guidance (Joreskog, 1993; Kline, 1998).

The inter-item correlation matrix was scrutinized to determine if it was suitable for exploratory factor analysis. Bartlett's test of sphericity resulted in a significant chi-square , χ^2 (253, n = 211) = 667.98 $p < .001$; thus, the hypothesis that the correlation matrix was an identity matrix was rejected. However, when the matrix was subjected to the KMO measure of sampling adequacy, the results were not as heartening (KMO = .614). This value is considered mediocre for factor analysis (Kaiser, 1974). Review of the eigenvalues and the scree test suggested a parsimonious three-factor solution might be possible. Therefore, three factors were extracted and orthogonally rotated using Varimax. The three factors accounted for 28.9% of the total variation of the 23 variables submitted to principal components analysis. The tentative names and loadings on the three factors retained for study are presented in Table 2. All listed variables loaded principally on the factor for which they are reported; in other words, no variable listed in Table 2 had a factor loading with an absolute value above .30 on any other factor. The simple structure of the factor solution was evident from the final rotation, since no variable loaded on more than one factor.

The first factor extracted consisted of 11 statements primarily related to libertarian views. Six of the items had been classified originally as antiauthoritarian and five items as equalitarian. The internal consistency of the items comprising the factor was alpha =.63. This is interesting, because despite the fact that 12 items were removed from the original scale, the homogeneity of the items was not diminished. While the content of items contributing to this factor seem to emphasize libertarian attitudes, there is another characteristic of these items that bears mention; all of these items were stated in the same direction. In other words, all of the items that loaded on the first factor, regardless of their original classification by Boehm (1968) as antiauthoritarian or equalitarian were stated in a positive direction such that agreement was indicative of libertarian attitudes and beliefs. Responses to these items were then reverse coded following the Kravitz et al. (1993) scoring guidelines to form an indication of legal authoritarian beliefs. Thus, high scores on the libertarian scale were reversed coded to indicate more legal authoritarian attitudes. The fact that none of the items loading on this scale required respondents to indicate disagreement to conform to libertarian views presents a potential measurement flaw (Crocker & Algina, 1986; Dawes, 1972; Haladyna, 1994).

The second factor consisted of 5 items. Individuals scoring high on this factor would endorse authoritarian views. The internal consistency of the factor was .57, which is fairly impressive given the small number of items contributing to it. All items loading on this factor originated from the original Boehm (1968) authoritarian scale. Parallel to the first factor items, all of these items were stated such that agreement with the statement would be associated with authoritarian or right wing views. This response consistency characteristic represents the same measurement threat described above.

The third factor, which consisted of only 4 items, seemed to solicit respondents' perceptions of police. This factor had an internal consistency coefficient of .51, and was comprised of two items originally classified as authoritarian and two equalitarian items. Table 3 provides descriptive statistics for the scales derived from the factor solution.

Correlation coefficients were computed between factor scores from the three factors and the responses from the four demographic questions shown in Table 1. Results indicated significant correlations between scores on Factor 1 and political orientation, in that subjects scoring higher on Factor 1 (higher legal authoritarian) tended to be more Republican in orientation, *r* (218) = .21, *p* = .002, more conservative, *r* (217) = .23, *p* < .001, favor the death penalty, *r* (217) = -.18, *p* = .007, and favor parental notification in the case of abortion, *r* (218) = -.13, *p* = .049. For Factor 2, higher scores (more authoritarian) were associated with being identified with the Republican party, *r* (216) = .19, *p* = .005 and favoring the death penalty, *r* (215) = -.14, *p* = .04. There were no significant correlations involving Factor 3. Correlations among the demographic variables confirmed the chi-square analysis above in that subjects identifying with the Republican party were more conservative, *r* (221) = .51, *p* < .001, and favored the death penalty, *r* (221) = -.20, *p* = .003. Participants who were more conservative in orientation favored parental notification, *r* (222) = -.20, *p* = .002.

TABLE 3 RLAQ23 Scale Descriptive Statistics

Factor Name	No. of Items	Mean	SD	Potential Range	Alpha
Libertarian	11	40.9	7.0	11-66	.63
Authoritarian	5	20.3	4.1	5-30	.57
Police	4	11.3	3.2	4-24	.51

The Kravitz, et al. (1993) scale modifications were quite successful in removing the ipsative nature of the scale scores, eliminating the dependency associated with the original forced-choice triad item design, and reducing error and missing data incidence. However, results of the current study suggest, just as Kravitz, et al. (1993) recommended, that further revision is necessary to create a psychometrically sound instrument. Several recommendations for future work appear warranted.

Additional items need to be created for each of the assessed dimensions to enhance reliability. There is currently too much error variance associated with each scale. Some authors (Crocker & Algina, 1986; Dawes, 1972; Haladyna, 1994) provide guidelines for writing and reviewing inventory items. Among the possible modifications that are consistent with these guidelines are the following: 1) a better balance of positively and negatively expressed statements within each subscale and across the instrument; 2) avoidance of statements containing universals such as "all," "always," "none," and "never"; 3) avoidance of indefinite qualifiers such as "only," "just," "merely," "many," "few," or "seldom"; 4) avoidance of use of negatives such as "not," "none," "never"; and 5) avoidance of "double-barreled" items. These modifications enhance reliability and assist in control for response sets. Further, by increasing the reliability of the scales, the potential for raising validity is enhanced.

An additional methodological consideration is appropriate samples for scale improvement research. Pilot instrument research is best conducted with a group comparable to that for which the test is designed; in this case, the best sample would be jury eligible adults. The use of homogeneous samples, such as college students, restricts the variability of the scores, which, in turn, restricts reliability. The Kravitz, et al. (1993) reliability differences observed from their first (using undergraduates) to their second study (using jury-eligible adults) demonstrated this point quite well. The current study supports the contention that reliability using college student samples underestimates scale reliabilities for the target population. The current study extends the work of Kravitz, et al. (1993) by again refuting the unidimensional structure of the RLAQ23, identifying three independent dimensional underlying responses, and providing preliminary evidence of content and construct validity.

REFERENCES

Arbuckle, J. (1995). *Amos users' guide*. Chicago, IL: SmallWaters Corporation.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238-246.

Boehm, V. R. (1968). Mr. Prejudice, Miss Sympathy, and the authoritarian personality: An application of psychological measuring techniques to the problem of jury bias. *Wisconsin Law Review, 1968*,734-750.

Cowan, C. L., Thompson, W.C., & Ellsworth, P. C. (1984). The effects of death qualification on jurors' predisposition to convict and on the quality of deliberations. *Law and Human Behavior, 8*, 53-79.

Crocker, L. M, and Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Harcourt Brace Jovanovich College Publishers.

Dawes, R. M. (1972). *Fundamentals of attitude measurement*. New York, NY: John Wiley.

Fahringer, H. (1993). "Mirror, mirror on the wall…": Body language, intuition, and the art of jury selection. *American Journal of Trial Advocacy, 17*, 197-204.

Fulero, S. M. & Penrod, S. D. (1990). The myths and realities of attorney jury selection folklore and scientific jury selection: What works? *Ohio Northern University Law Review, 17*, 229-253.

Haladyna, T. M. (1994*). Developing and validating multiple-choice test items*. Hillsdale, NJ: Erlbaum.

Jeans, J. W. (1993). *Trial advocacy* (2nd ed.). St. Paul, MN: West Publishing.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39*, 31-36.

Joreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Lang (Eds.), *Testing Structural Equation Models* (pp. 294-316). Newbury Park, CA: Sage.

Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York, NY: The Guilford Press.

Kravitz, D. A., Cutler, B., L., & Brock, P. (1993). Reliability and validity of the original and revised Legal Attitudes Questionnaire. *Law and Human Behavior, 17*, 661-677.

Loehlin, J. C. (1992). *Latent variable models: An introduction to factor, path, and structural analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Publishing.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 103*, 391-410.

Moran, G., & Comfort, J. C. (1982). Scientific juror selection: Sex as a moderator of demographic and personality predictors of impaneled felony juror behavior. *Journal of Personality and Social Psychology, 43*, 1052-1063.

Narby, D. J., Cutler, B. L., & Moran, G. (1993). A meta-analysis of the association between authoritarianism and jurors' perceptions of defendant culpability. *Journal of Applied Psychology, 78*, 34-42.

Schutte, J. W. (1996). What do I need a psychologist for? *Trial Diplomacy Journal, 19*, 269-273.

Singer, A. (1995). Six rules for picking a successful jury. *The Practical Litigator, 6*, 15-18.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*, 1-10.

Vinson, D. E. (1993). *Jury persuasion: Psychological strategies and trial techniques.* Englewood Cliffs, NJ: Prentice Hall.

Author Note: We thank Robin Anderson, Joseph Basgier, Cindy Cull, Jill Maliszewski, Sean Siegel, Tanya Wade, Erin Winters, Latisha Gardner, Mathre Harper, Jill Kovaly, Ashlea Philipps, Jennifer Trager, Lynn Vogel, and John Willse who acted as research assistants.

Life after College: Psychology Students' Perceptions of Salary, Business Hiring Criteria, and Graduate Admission Criteria

Deborah S. Briihl
Valdosta State University

Ninety undergraduate psychology majors at a mid-size public southeastern regional university were asked to rate the importance of various characteristics for obtaining a job or getting into graduate school. They were also asked (1) what they perceived to be the minimum GPA and GRE scores for individuals getting into graduate programs; (2) the average GPA and GRE scores of individuals accepted into graduate programs; and (3) estimated starting salaries for individuals with a bachelor's, master's, or doctoral degree in psychology. Findings revealed that students are aware of most of the top criteria used by businesses and graduate schools and the minimum GPA and GRE scores; however, they were unaware of the relative importance of other criteria. Students tended to underestimate the average GPA and GRE scores of individuals accepted into graduate programs and to overestimate starting salaries.

One of the biggest questions that undergraduate students must face is what they will do after college. Should they seek a job or try to go to graduate school? Whatever path they choose, students need to be aware of the factors necessary for success in gaining either employment or entrance into graduate school. In this study, psychology undergraduate students were asked what they perceive to be the business hiring and graduate school admission criteria. These responses were compared to the actual criteria used by both businesses and graduate schools.

Previous research has summarized the average GPA and GRE scores for psychology students in both master's and doctoral level programs (Norcross, Hanych, & Terranova, 1996), second-order criteria used by graduate schools (Keith-Spiegel, Tabachnick, & Spiegel, 1994; Landrum, Jeglum, & Cashin, 1994), and the usefulness of a master's degree for gaining admission into a doctoral program (Bonifazi, Crespy, & Rieker, 1997). The two most important objective criteria for gaining admission to graduate school are GPA and GRE. In order to get into a master's program, the average minimum GPA is 2.86 and GRE score of 984. The average scores of students accepted into a master's program are 3.27 GPA and 1033 GRE. For a doctoral program, the average minimum GPA

Author info: Correspondence should be sent to: Dr. Deborah Briihl, Dept. of Psychology & Counseling, Valdosta State U., Valdosta, GA 31698.

North American Journal of Psychology, 2001, Vol. 3, No. 2, 321-330.
© NAJP

is 3.09 and GRE score of 1090. The average scores of students accepted into doctoral programs are 3.50 GPA and 1206 GRE (Norcross, et al., 1996). The nonobjective requirements for acceptance into doctoral or master's programs are the personal statement, letters of recommendation, and the interview. Doctoral programs are more likely to rate research as highly important than are master's programs (Keith-Spiegel et al., 1994; Landrum, et al., 1994).

Other studies have asked businesses what criteria are used for hiring decisions (NACE, 1995a, 1997, 1999). The most important factors are (in descending order) communication skills, past experience, motivation, teamwork, GPA, interpersonal skills, technical and analytical skills. While GPA was rated of medium importance, 63% of businesses that use GPA as a prescreening qualifier used 3.0 as their cutoff score (NACE, 1995b). GPA can also impact starting salary. According to the *Occupational Outlook Handbook* (U.S. Bureau of Labor Statistics, 2000-01), the starting salary of someone with a bachelor's in psychology was higher if the student had superior academic records. While the interview is not among the items listed, employers have also discussed the importance of first impressions as well as stressing communication skills - both of which are conveyed in the interview (Allen, 1997; Perry & Goldberg, 1998).

Are undergraduate psychology students aware of what is important for success in obtaining a job or getting into graduate school? Few studies have addressed this issue (e.g. Cashin & Landrum, 1991). In the present study I asked students their knowledge about the process. Students were asked what they thought the average starting salary was for someone with a bachelor's, master's, and doctorate in Psychology. They were also asked what they thought the minimum GPA and GRE scores were for master's and doctoral programs and the average GPA and GRE of students that have been accepted into master's and doctoral programs. Students then judged the importance of various criteria when seeking a job (with a bachelor's degree) and when trying to get into graduate school.

METHOD

Participants

Ninety undergraduate psychology majors from a regional southeastern university completed the questionnaire. The majority of the participants were women (79%) and seniors in college (64%). When students were asked to indicate their future plans, 75% stated they planned on going to graduate school at some time in the future.

Materials and Procedure

Students were asked what they believed a person starting out with a bachelor's, a master's, and a doctorate in Psychology would earn. They were then asked what they thought the minimum GPA and GRE scores were for someone to get into a master's and a doctoral program in psychology and what the average GPA and GRE score were for people who had gotten into a master's and a doctoral program. Finally, they were given a list of factors and asked to judge the important of each. For businesses, students rated the importance of 14 factors that a business might use when they are looking to hire someone with a bachelor's degree in psychology. For graduate school, students rated the importance of 15 factors that a typical graduate program might use when it makes admissions decisions. Students were told to use the following key: 4=one of the main factors, 3=important, but not one of the main criteria used, 2=minimally important, 1=not important.

RESULTS

Table 1 lists the student responses to the salary questions, and Table 2 lists the responses to the GPA and GRE questions. The data are split between Senior and Sophomore/Junior responses. Since there were few sophomore responses (10% of all responses) and students do not declare their major until the middle of their sophomore year, the data for sophomores and juniors were combined. Students tended to overestimate starting salaries, with seniors more realistic about starting salaries than

TABLE 1 Undergraduate psychology students' perceptions of average starting salaries for an individual with a Bachelor's, a Master's, and a Doctorate in psychology.

	Sophomores and Juniors	Seniors	Actual
Bachelor's	$25,307 (5,676)	$23,000 (4,861)	$22,000
D/k	8%	3%	
Master's	$35,666 (8,230)	$32,557 (7,306)	mid $20-mid $30,000
D/k	6%	10%	
Doctorate	$50,107 (15,504)	$44,865 (13,357)	low $30-mid $40,000
D/k	12%	10%	

Note: Standard deviations are in parentheses. D/k stands for the percent of students who did not respond to this question or wrote "don't know." Included for comparison are the actual starting salary figures for individuals with a Bachelor's (Traub & Allen, 1996), Master's degree (APA, 1999a), and Doctorate in psychology (APA, 1999b).

sophomores and juniors, although approximately 10% wrote "don't know" or left the question blank. Students were fairly accurate about the minimum GPA requirements needed to get into both a master's and a doctoral program. Sophomores and Juniors were more likely to overestimate GRE scores as compared to Seniors. Both Seniors and Juniors/Sophomores underestimated the median GPA of students who have been accepted for both the master's and the doctorate programs. These numbers do not include the relatively small percent (2 - 10%) of the students whose responses to the minimum and average GPA questions were "don't know" or left the question blank.

TABLE 2 Undergraduate psychology students' perceptions of the minimum GPA and GRE scores needed to get into Master's programs and Doctoral programs and the average score of students in Master's and Doctoral programs.

		Sophomores & Juniors	Seniors	Actual
Master's				
	Minimum GPA	2.99 (.23)	2.89 (.26)	2.86
	D/k	3%	2%	
	Minimum GRE	1093 (135)	995 (130)	984
	D/k	63%	17%	
	Average GPA	3.08 (.31)	3.14 (.22)	3.27
	D/k	9%	5%	
	Average GRE	1112 (108)	1065 (133)	1033
	D/k	63%	27%	
Doctoral				
	Minimum GPA	3.21 (.35)	3.29 (.29)	3.09
	D/k	3%	2%	
	Minimum GRE	1166 (81)	1140 (193)	1090
	D/k	63%	22%	
	Average GPA	3.35 (.31)	3.38 (.24)	3.50
	D/k	9%	8%	
	Average GRE	1212 (102)	1225 (159)	1206
	D/k	71%	29%	

Note: Standard deviations are in parentheses. D/k stands for the percent of students who did not respond to this question or wrote "don't know." Included for comparison are the actual minimum and average scores of students in Master's and Doctoral programs (Norcross, et al., 1996).

One of the more surprising findings was the question about GRE scores. On average, 63% of the sophomores and juniors wrote "don't know" or left the question blank. Three juniors wrote down numbers that were impossible to obtain from the GRE test (2000, 34, 35). More seniors ventured a guess; however, between 17-30% of the responses were "don't know" or left blank. Only one senior wrote down an impossible GRE score (34). This finding was quite surprising considering that 75% of the respondents stated that they were planning on going to graduate school.

Table 3 presents the criteria ranked according to their perceived importance. Students accurately rated past experience and communication skills as the top factors that employers look for when they are hiring someone. Students underestimated the importance of the interview, letters of reference, and GPA in getting a job. Students accurately indicated that important criteria for acceptance into graduate school were GPA and GRE scores. However, they viewed an internship to be more important than research experience, letters of reference, and the interview. Students also underestimated the importance of the personal statement. Most students knew that one does not need a master's degree before applying to most doctoral programs.

TABLE 3 Undergraduate psychology students' average responses to the importance of the following criteria to a business making hiring decisions and a graduate program making admissions decisions.

Business		Graduate School	
3.26	Communication Skills (listening, speaking, writing)	3.32	GPA (for coursework)
3.23	Past Experience	3.19	Breadth of course taken (includes minors)
3.21	Personal Skills (time management, motivation)	3.11	GRE scores
3.08	Information Gathering Skills (logical thinking, planning and carrying out projects)	3.06	GPA (overall)
2.92	Interview	3.01	Internship
2.90	Breadth of course taken (includes minors)	2.96	Information Gathering Skills (logical thinking, planning and carrying out projects)
2.87	Computer Skills	2.94	Communication Skills (listening, speaking, writing)
2.72	Letters of Reference	2.86	Letters of Reference
2.59	Social Skills (ability to work with others)	2.83	Research
2.49	GPA (overall)	2.82	Interview
2.46	GPA (for coursework)	2.80	Social Skills (ability to work with others)
2.36	Cover letter/personal statement	2.51	Computer Skills
2.28	Research	2.48	Personal Statement
2.15	Extracurricular activities	2.26	Extracurricular Activities
		1.64	Importance of a Master's degree to get into a Doctoral program

Note: The numbers are based on the 4 point scale used for the survey with 4=one of the main factors, 3=important, but not one of the main criteria used, 2=minimally important, 1=not important.

DISCUSSION

This study attempted to analyze students' perceptions of the job and graduate school criteria selection as well as starting salaries. A number of useful findings emerged. First, students (particularly Sophomores and Juniors) were overly optimistic about the amount of a starting salary. While the average starting salary has increased over the past few years, a person with a bachelor's degree in psychology still tends to have one of the lower starting salaries ($19,500-$22,300) as compared to business ($27 - 33,000), computer science/engineering ($32 - 41,000), health sciences ($25 - 49,000), or other sciences ($24 - 29,000) and Humanities ($23 - 25,000) (Traub & Allen, 1996; U.S. Bureau of Labor Statistics, 2000-01). This trend also is true at both the master's and doctoral level. According to the APA (1999a, 1999b), the starting salary for someone with a master's ranged from the mid $20's to mid $30's and for someone with a doctorate, between the low $30's and mid $40's. Individuals with an M.B.A. can start at $42,000 and individuals with a doctorate in the other sciences (biology, chemistry, physics) can start at $45-60,000 (U.S. Bureau of Labor Statistics, 2000-01). This difference between psychology and the other sciences also can be seen in the academic setting. Faculty members teaching in those areas that have high-paying nonacademic jobs tend to be the ones making higher salaries.

Seniors were fairly accurate about the objective criteria necessary to get into graduate school, both at the master's and doctoral levels, although both groups underestimated GPA scores. Students need to be made aware of the fact that average GPA scores are higher than the minimum requirements and that merely having the minimum GPA does not guarantee acceptance into a graduate program, particularly given the high number of applicants to doctoral programs. An interesting finding was the lack of knowledge about GRE scores. While Seniors were more knowledgeable than Sophomores and Juniors about the necessary scores, a large percentage of the Seniors were still unable to guess. The results are even more surprising considering that the majority of the students in this study stated that graduate school was part of their future plans. Students typically take the GRE at the beginning of their senior year. However, they should have started preparing for it a few months before they take the exam.

As mentioned earlier, the main criteria to get into graduate school are GPA, GRE, personal statement, letters of recommendation, and the interview (Bonifazi, et al., 1997; Keith-Spiegel, et al.,1994; Landrum, et al, 1994; Norcross, et al., 1996). Students accurately indicated that important criteria for acceptance into graduate school were GPA and GRE scores. However, they viewed an internship to be more important than research experience, letters of reference, and the personal statement.

Students also underestimated the importance of the interview. It should be noted, however, that the graduate program at the university where the study took place does not use the interview as part of the application process.

The main criteria for obtaining a job are communication skills, past experience, motivation, teamwork, GPA, interpersonal skills, technical and analytical skills (NACE, 1995a, 1997, 1999). Students accurately rated past experience and communication skills as the top factors that employers look for when they are hiring someone. Students underestimated the importance of the interview, letters of reference, and GPA in getting a job. Students overestimated the importance of information gathering skills and underestimated the importance of social skills.

These results suggest that, while students know what primary factors are necessary to obtain a job or get into graduate school, they are not as familiar with second-order criteria that are commonly used. According to Keith-Spiegel, et al. (1994), doctoral programs receive a large number of applicants and first order criteria are not enough to narrow the field to the number of slots available. Subsequently, they must rely heavily on these second-order criteria to make final acceptance decisions. It is important for students to realize that they need to focus their time and resources not just on grades, but also on other areas, particularly research, if they wish to get into a graduate program. Students need to be aware that the minimum requirements needed to get into a graduate program do not guarantee acceptance. Students knew the minimum requirements, but they underestimated the average GPA and GRE scores of students in master's and doctoral programs.

The same holds true for obtaining a job. Many students were not aware of the fact that both graduate programs and future employers look at GPA when they consider applicants. In addition, students also need to promote the skills that they obtain with an undergraduate degree in psychology. Many of the skills that employers are seeking are those that students obtain in a bachelor's program – such as communication, time management, social, and information processing skills. According to the APA Research Office (1999c), between 1996 and 1997 over 70,000 bachelor's degrees were awarded in psychology. During that same time period, between 10,000 and 20,000 students obtained master's degrees in psychology and less than 10,000 obtained a Ph.D. It is obvious that the majority of students who get a bachelor's degree in psychology do not go to graduate school. As advisors, it is important for us to not only make sure that students know the criteria for acceptance into graduate programs, but also the criteria necessary to obtain a job.

How can we decrease the amount of misinformation that our students seem to have? As instructors, we can offer career education courses, include career projects in our courses that require students to obtain this information, or create handouts or web sites with this information. Many institutions have career centers on campus. As advisors, we can encourage our students to visit the center or invite individuals from the career center to our classes or to psychology club meetings. In addition, we, as advisors, need to become more knowledgeable about these requirements. Many institutions have little formal advisor training beyond a basic understanding of the coursework that students need to complete. Perhaps more institutions need to develop a better way of training advisors, either at the institution level or in graduate programs. These interventions would provide more accurate information to students, enabling them to make better, informed decisions about their college and future career plans.

REFERENCES

Allen, C. (1997). Interviews. *Spotlight, 19,* 7.

APA Research Office (1999a). 1996 employment survey: Psychology graduates with Master's, Specialist's, and Related Degrees [On-line]. Available: http://research.apa.org/mas5.html [1999, November].

APA Research Office (1999b). Preliminary data: 1997 Doctorate employment survey [On-line]. Available: http://research.apa.org/doc06.html [1999, Nov.].

APA Research Office (1999c). Data on education and employment: Number of psychology degrees conferred by level of degree: 1996-1997 [On-line]. Available: http:// http://research.apa.org/gen2.html [1999, November].

Bonifazi, D.Z., Crespy, S.D., & Rieker, P. (1997). Value of a master's degree for gaining admission to doctoral programs in psychology. *Teaching of Psychology, 24,* 176-182.

Cashin, J.R., & Landrum, R.E. (1991). Undergraduate students' perceptions of graduate admissions criteria in psychology. *Psychological Reports, 69,* 1107-1110.

Keith-Spiegel, P., Tabachnick, B.G. & Spiegel, G.B. (1994). When demand exceeds supply: Second-order criteria used by graduate school selection. *Teaching of Psychology, 21,* 79-81.

Landrum, R.E., Jeglum, E.B., & Cashin, J.R. (1994). The decision-making processes of graduate admissions committees in psychology. *Journal of Social Behavior and Personality, 9,* 239-248.

National Association of Colleges and Employers (1995a). *Spotlight, 17,* 1-2.

National Association of Colleges and Employers (1995b). 1994 college relations and recruitment survey: A measure of the HR recruitment function. *Journal of Career Planning and Employment, 55,* 1-12.

National Association of Colleges and Employers (1997). 1997 Job Outlook Survey [On-line]. Compiled by Research Office, APA. Available: http://research.apa.org/bac2.html [1999, November].

National Association of Colleges and Employers (1999). What wows employers? *Job Outlook 99*, 6-7.

Norcross, J.C., Hanych, J.M., & Terranova, R.D. (1996). Graduate study in psychology: 1992-1993. *American Psychologist, 51*, 631-643.

Perry, A., & Goldberg, C. (1998). Who gets hired? Interviewing skills are prehire variable. *Journal of Career Planning and Employment, 58*, 47-50.

Traub, D., & Allen, C. (1996). Hiring up, starting salaries higher for new college grads. *Salary Survey* [On-line]. Available: http://www.jobweb.org/Pubs/Pr/pr090996.htm [1999, November]

U.S. Bureau of Labor Statistics (2000-01). *Occupational Outlook Handbook* Washington, D.C.: U.S. Government Printing Office.

The Effect of Politeness and Grammar on User Perceptions of Electronic Mail

Sherri L. Jessmer
North Rose, New York

David Anderson
Allegheny College

The present study examined the effect of the nature (polite vs. impolite) and the format (grammatical vs. ungrammatical) of an e-mail message on how the sender of that message was perceived by the recipient. Nine attributes of the sender were examined: status, gender, competency, status in relation to the recipient, how much the recipient would like to work on a regular basis with the sender, friendliness, how likable the sender was, if the message had been edited, and if the recipient thought that the sender cared how the recipient was going to react to the message. In a sample of 112 undergraduates, senders of polite grammatical messages were viewed the most positively; impolite messages were seen as being authored by high status males; and polite messages were seen as originating with competent females. No significant differences were found between the two data collection methods that were used: e-mail vs. paper presentation of the questionnaire.

Electronic mail (e-mail) is just the most recent in a long line of alternatives to face-to-face communication. The personal, hand-written letters of the nineteenth century were replaced by the telephone call of the twentieth-century; and e-mail is replacing them both in the twenty-first century (NUA Internet Surveys, 1998). But even if it is not true that "the medium is the message," we do know that the e-mail medium is different in many ways from these other means of communication (Adkins & Brashers, 1995; Lea & Spears, 1992). The present research examined how the nature and format of an e-mail message might affect how a recipient perceives the sender of that message.

Carnevale and Probst (1997) pointed out that text based communication makes interaction more difficult because it is less "rich" than other forms of communication. For example, a number of researchers have noted the lack of social and status cues in e-mail (Carnevale & Probst, 1997; Garton & Wellman, 1995; McCormick &

Author info: Correspondence should be sent to: David Anderson, Dept. of Psychology, Allegheny College, Meadville, PA 16335. E-mail: danders@allegheny.edu

North American Journal of Psychology, 2001, Vol. 3, No. 2, 331-346.

© NAJP

BEST COPY AVAILABLE

McCormick, 1992) while Sproull and Kiesler (1986) suggested that although people make use of both static (e.g., appearance) and dynamic (e.g., nodding) cues as they communicate, dynamic cues are completely missing from e-mail. Static cues, on the other hand, are present, although they may appear in a different way than in face-to-face communication.

In an extensive review of the literature on language and persuasion, Burgoon and Miller (1985) noted that language is a rule-governed system and that we all develop expectations and norms concerning usage in particular circumstances. (Also see Adkins & Brashers, 1995 and Holmes, 1995.) But what happens when those expectations and norms are violated? For example, how do we react when the language in an e-mail message is ungrammatical or impolite?

Experiments in impression formation suggest that we often make use of social schemas to make judgements about the characteristics of a role, an event, a person, or a group (Fisk & Taylor, 1991). Some of the most pervasive social schemas concern gender differences, i.e., expectations about what is appropriate behavior for males and females. Eagly (1987) has suggested social norms provide a way to understand the source of these expectations. According to Eagly, social norms are shared expectations about appropriate behavior. A major assumption of the social-role interpretation of sex differences is that the perception of women as especially communal and men as especially agentic stems from the differing specific roles that women and men occupy in society. She further notes these pervasive differences in gender roles lead to differences in social status and power.

Recently, an interesting analysis of men's and women's speech and its relationship to status and power was published (Holmes, 1995). Holmes not only found differences in politeness patterns between men and women, but she noted that the more polite tone of women's speech is often associated with submissive social roles.

Since context cues often act as reminders of social norms, their absence can lead to reduced attention to the behavioral norms that usually govern interpersonal behavior. For example, Lea and Spears (1991) noted that as an interaction becomes depersonalized, communication also becomes less personal. They suggest that as we become less aware of our audience, we become secure in our anonymity and hence less bound by social norms. Several researchers (Carnevale & Probst, 1997; Garton & Wellman, 1995; Weisband & Reining, 1995) suggest that employees who communicate with their supervisors by means of e-mail could begin to lose the sense of the social status of their supervisor. This in turn could lead to miscommunication that could, in extreme cases, lead to hostility and flaming (Welch, 1997). Chapman (1995) added that when people are completely physically removed from each other they feel safe in verbally assaulting the recipient of their messages.

Contextual cues also serve to help us make assumptions about the sender of a message, including gender, age, and race. However, as Sproull & Kiesler (1986) pointed out, when there are few such cues, senders and receivers may assume that they possess the same qualities since they are unable to determine the race or gender of the other.

When status cues are reduced in e-mail messages, the recipient must decide the status of the sender based on other information. Assumptions about the status of the individual sending a message can influence how the recipient interprets the message. As Lee (1996) noted, there is a large body of social science research that indicates that status is a key component to how persuasive a person is perceived to be. High status individuals are usually considered more competent and more knowledgeable than low status people (Lee, 1996).

Stahelski and Paynton (1995) maintained that people of high status are perceived as using what they called a "strong strategy" composed of high authority, sanctions, reasoning, and a display of power. On the other hand, people of lower status are perceived as using a weak strategy composed of friendliness, assertiveness, coalition, influence tactics, and less control of resources. But how can one determine status in electronic communication when so few cues are available?

Sherwood (1998), the creator of a website on how to write effective e-mail, suggested that one of the biggest cues people use to determine status is by the way the writer of the message uses language. The accurate use of large words, but not in over abundance, can imply high status, for example. Poor grammar and misspellings can imply that the sender is uneducated and hence of low status.

In the same vein, Glassberg, Kettinger, and Logan (1996) pointed out that e-mail messages tend to de-emphasize professional titles, such as "Dr." or "President of." User names usually do not include clues to professional status. This is in contrast to status cues provided by a letterhead or by dress in face-to-face meetings.

The role of language in determining gender has been extensively examined. Unfortunately, most of the research has focused on verbal or written communication, with little work on e-mail. The work that has been done, suggests that men use talking for attention purposes, as a way of getting and keeping attention, while women use talk for relationship purposes and to maintain connections (Rodino, 1997). Women tend to make more use of qualifiers, exaggerations, extreme politeness, and tag questions while put-downs, strong assertions, and sarcasm characterize male language. These conclusions were based on the analysis of conversations from an Internet Relay Chat channel (Rodino, 1997).

In one of the few attempts to verify these assertions in electronic communication, Witmer and Katzman (1997) examined messages posted to public newsgroups. They found that females used graphical accents more than men did, thus suggesting more emotional tone in their messages.

The final question in this study concerns the use of e-mail itself to gather data. As more and more research is being done electronically, it becomes more important to determine what, if any, differences there are based on the mode of presentation of a questionnaire. To date the research seems to suggest that electronic responses tend to be more extreme, more self-disclosing, provide less socially desirable answers (Kiesler & Sproull, 1986; Sproull, 1986), and are more self-absorbed (Kiesler & Sproull, 1986).

As contextual cues are reduced in a communication environment such as e-mail, the recipient of a message must make assumptions about the sender of the message based on the information that is available within the message itself. The literature suggests that assumptions about gender (Allen, 1995; Sproull & Kiesler, 1986) and status and competency (Lee, 1996; Sherwood, 1998; Stahelski & Paynton, 1995) are often determining factors in how people respond to e-mail. In an attempt to tease out some of the specific components of e-mail that might effect these assumptions, the present study examined the influence of the nature of a request (polite or impolite) and format of a request (grammatical or ungrammatical) on a user's perception of the status, gender, and competency of the sender. In addition, the researchers explored two methods of data collection, paper-and-pencil versus e-mail presentations.

Based on the literature, four hypotheses were proposed for examination.

H1: It was hypothesized that grammatical messages would be perceived to have been written by a person of higher status than ungrammatical messages.

H2: Since grammatical messages would be seen as originating with higher status individuals, it was assumed that people of higher status would also be seen as more competent.

H3: It was proposed that polite message authors would more often be perceived as being female.

H4: Finally, concerning socially desirable responses, it was proposed that in the results obtained via e-mail we would see an exaggeration of the findings noted above. For example, if the perception of gender is significantly affected by the nature of the message, then this result would be larger in the group who responded to the questionnaire via e-mail compared to those who completed a paper version.

METHOD

Participants

One hundred twelve undergraduates (80 females and 32 males, ages 18-21, 95%+ white) volunteered to participate in the study. Fifty-seven (42 females and 15 males) completed the paper version of the questionnaire and 55 (38 females and 17 males) the e-mail version. Students indicated their preference for paper or e-mail versions of the questionnaire on a sign-up sheet that was distributed to all sections of introductory psychology (none taught by the authors). All students received extra credit in an introductory psychology class for participation.

Materials

Four messages that varied in the nature of the request (polite or impolite) and the format of the request (grammatical or ungrammatical) were created. This produced messages that were polite and grammatical; impolite and grammatical; polite and ungrammatical; and impolite and ungrammatical. Messages were one paragraph in length and addressed the need to limit the use of copiers and printers to business related activities.

The polite, grammatical message read as follows: "This is just a note to request that the main copiers and printers not be used for personal copies. Please make personal copies using only the copier in the break room. Unfortunately, there is a small charge for copies on this machine but in order to reduce costs within the office we need to try to maintain this policy. We are sorry to cause any inconvenience by the enforcement of these policies. I hope that this is the last time this issue will need to be addressed."

The impolite messages threatened "strict repercussions" and noted that "THIS IS THE LAST TIME THIS ISSUE WILL BE BROUGHT UP." Ungrammatical messages included misspellings, typing errors ("PLEase"), missing punctuation, and fractured grammar (e.g., "FEE 10 cents per copy" and "iT will not be addressed again").

Design and Procedure

To maximize the number of responses to each message, a within-subjects design was used. That is, all participants were asked to read and respond to all four of the messages. Participants were asked to imagine that they had received each of the four e-mail messages from an anonymous sender at their place of work. They were instructed to carefully read each one and then respond to 12 seven-point Likert scale items.

Approximately one-half of the participants completed a paper-and-pencil version of the questionnaire during several group out-of-class testing sessions. The other half received the same questionnaire via e-mail and returned it electronically to the experimenters. The result was a 2 x 2 x 2 design: first two factors within-subjects, third factor between-subjects.

In addition to three items designed as manipulation checks, nine seven-point Likert scales asked the reader to judge several qualities of the message. These included: power/status of the sender, power/status of the sender in relation to the recipient, gender of the sender, competency of the sender, friendliness of the sender, likeability of the sender, the editing of the message, overall reaction to the message, and the extent to which the recipient would want to work with the sender on a regular basis.

Upon completion of the paper questionnaires, participants were provided with a debriefing statement outlining the purpose of the study. E-mail participants were provided with the same statement by return e-mail. Before the data were analyzed all identifying information was removed from the e-mail questionnaires.

RESULTS

Manipulation Checks

Manipulation checks were examined to ensure that the levels of the independent variable were perceived as they were intended. Participants were asked about the grammatical structure of each message, the politeness of the message, and the format of the message. Messages were compared across the entire sample (grammatical versus ungrammatical; polite versus impolite; and well formatted versus poorly formatted) as well as within each presentation method (e.g., paper presentation: grammatical versus ungrammatical).

Grammatical messages ($M = 5.24$, $SD = 1.55$) were reported to be more grammatical than ungrammatical messages ($M = 1.68$, $SD = 1.20$), $t_{446} = 27.10$, $p < .01$. In addition, the polite messages ($M = 5.00$, $SD = 1.74$) were viewed as more polite than impolite messages ($M = 2.53$, $SD = 1.31$), $t_{446} = 16.89$, $p < .01$.

These manipulations were effective in both manners of presentation. In the e-mail responses, grammatical messages ($M = 5.12$, $SD = 1.77$) were seen as more grammatical than ungrammatical messages ($M = 1.96$, $SD = 1.46$), $t_{218} = 14.41$, $p < .01$. Also, the e-mail polite messages ($M = 5.11$, $SD = 1.88$) were perceived as more polite than the impolite messages ($M = 2.56$, $SD = 1.44$), $t_{218} = 11.27$, $p < .01$.

In the paper responses, the polite messages ($M = 4.89$, $SD = 1.60$) were agreed to be more polite than the impolite messages ($M = 2.50$, $SD = 1.17$), $t_{226} = 12.80$, $p < .01$. In addition, the grammatical messages in the paper presentation ($M = 5.35$, $SD = 1.28$) were seen as more grammatical than the ungrammatical messages ($M = 1.42$, $SD = 0.82$), $t_{226} = 27.57, p < .01$.

TABLE 1 Means (and SDs) for Perceived Attributes of the Sender Of an E-mail Message as a Function of Message Type

	Message Type[a]			
Attribute	Polite	Impolite	Grammatical	Ungrammatical
Power	4.52 (1.48)	4.95 (1.32)	5.67 (1.19)	3.80 (1.58)
Gender*	3.82 (1.24)	4.50 (1.33)	3.95 (1.33)	4.37 (1.25)
Competency	4.27 (1.13)	3.94 (1.24)	5.57 (1.08)	2.64 (1.28)

Note: Values represent participants' agreement with statements that addressed each attribute on a 7-point scale (1 = very strongly disagree and 7 = very strongly agree).

[a] $N = 224$ for all message types.

* Represents agreement that the sender was male.

Perceived Attributes of the Sender

The three perceived attributes of the sender, status (power in the company) gender, and competency, were examined independently using 2 x 2 x 2 one-between (e-mail/paper participation), two-within analysis of variance (polite/impolite and grammatical/ungrammatical).

Status

No significant effects were found in the status of the sender of the messages as a function of the method of questionnaire administration ($F_{1,110} = 1.14$, $MSE = 3.27$). On the other hand, impolite messages were seen as being written by a person in a position of power ($F_{1,110} = 20.66$, $p < .01$, $MSE = 0.99$) while grammatical messages were seen as originating with individuals with more power ($F_{1,110} = 157.05$, $p < .01$, $MSE = 2.49$). The means used in these comparisons can be found in Table 1. No significant interactions were found between the politeness, grammar, and method of questionnaire administration.

Gender

There were no significant differences found on the perceived gender of the sender as a function of the manner of presentation ($F_{1,110} = 1.97$, $MSE = 3.33$). However, impolite messages were seen as being written by a male ($F_{1,110} = 42.05$, $p < .05$, $MSE = 1.22$) as were ungrammatical messages ($F = 15.62$, $p < .05$, $MSE = 1.30$). See Table 1 for details.

The only significant interaction concerning gender was found between the nature of the request and the format of the request ($F_{1,110} = 29.85$, $p < .05$, $MSE = 0.88$). Ungrammatical messages were seen as being authored by males whether they were polite ($M = 4.28$, $SD = 1.32$) or impolite ($M = 4.47$, $SD = 1.70$). However, grammatical messages were viewed as coming from males only if they were impolite ($M = 4.53$, $SD = 1.48$). That is, only with polite grammatical messages did the participants view the e-mail as coming from a female ($M = 3.37$, $SD = 1.16$).

TABLE 2 Means (and SDs) for Additional Attributes of the Sender Of an E-mail Message as a Function of Message Type

	Message Type			
Attribute	Polite	Impolite	Grammatical	Ungrammatical
Relative Power	4.75 (1.40)	5.23 (1.21)	5.71 (1.15)	4.27 (1.45)
Work With	4.00 (1.30)	2.85 (1.40)	4.19 (1.41)	2.66 (1.29)
Friendliness	5.02 (1.08)	3.34 (1.17)	4.40 (1.13)	3.96 (1.12)
Likable	4.84 (1.08)	3.39 (1.16)	4.38 (1.13)	3.84 (1.11)
Edit	3.75 (1.06)	3.36 (1.30)	5.66 (1.34)	1.45 (1.02)
Reaction	4.59 (1.39)	3.48 (1.85)	4.99 (1.65)	3.07 (1.62)

Note: Values represent participants' agreement with statements that addressed each attribute on a 7-point scale (1 = *very strongly disagree* and 7 = *very strongly agree*).
[a] $N = 224$ for all message types.

Competency of Sender

No significant differences were found on competency of the sender as a function of the manner of participation ($F_{1,110} = 0.61$, $p > .05$).

Higher competency ratings were given when the e-mail was polite ($F_{1,110} = 8.46$, $p < .05$, $MSE = 1.49$) and grammatical ($F_{1,110} = 509.92$, $p < .05$, MSE = 1.88). See Table 1 for means. None of the interactions involving the competency of the sender were significant.

Analysis of Additional Attributes

In addition to the main attributes noted above, several other attributes of the sender as perceived by the recipient were examined. A series of 2 x 2 x 2 one-between (e-mail/paper participation), two-within analyses of variance (polite/impolite and grammatical/ungrammatical) were done on these six attributes.

No main effects or interactions involving the method of presentation were significant for any of the six attributes. However, the same pattern emerged with the main effects and interactions of all six when message type and format were examined. 1) Grammatical messages were seen more positively; 2) Polite messages were rated more positively; and 3) polite/grammatical messages were perceived more positively (with one exception) than the other combinations. See Table 2 and Table 3 for the means and standard deviations and Table 4 for statistical results.

TABLE 3 Means (and SDs) for Additional Attributes of the Sender of an E-mail Message as a Function of the Interaction of the Format and the Nature of the Message.

	Message Type[a]			
	Polite		Impolite	
Attributes	Grammatical	Ungrammatical	Grammatical	Ungrammatical
Relative Power	5.34 (1.22)	4.16 (1.56)	6.08 (1.08)	4.38 (1.32)
Work With	5.38 (1.29)	2.63 (1.31)	3.01 (1.51)	2.69 (1.28)
Friendliness	5.74 (1.08)	4.31 (1.07)	3.06 (1.18)	3.62 (1.16)
Likable	5.53 (1.07)	4.14 (1.10)	3.23 (1.20)	3.54 (1.13)
Edit	6.04 (1.04)	1.45 (1.09)	5.28 (1.58)	1.44 (0.94)
Reaction	5.96 (1.07)	3.21 (1.66)	4.02 (2.07)	2.93 (1.59)

Note: Values represent participants' agreement with statements that addressed each attribute on a 7-point scale (1 = *very strongly disagree* and 7 = *very strongly agree*).
[a] *N* = 224 for all message types.

Participants were more likely to agree that the sender was more powerful (Relative status: "The sender of this message has more power than the recipient") when the message was polite or when the message was grammatical. And although polite/grammatical messages were seen originating from a more powerful sender when compared to polite/ungrammatical and impolite/ungrammatical e-mail, the most powerful senders were seen in messages that were grammatical but impolite. This complex interaction is the exception to the general findings noted above.

The attribute of "work basis" addressed the extent to which the recipient of the message thought that they would like to work with the sender on a regular basis. And as might be expected, recipients were more likely to want to work with the senders of polite messages and with the senders of grammatical messages. And finally, the interaction pattern noted above was seen with work basis. That is, messages that were both polite and grammatical elicited much more positive ratings that those of any other combination. See Tables 2 and 3 for means and standard deviations.

TABLE 4 Analysis of Variance F values (and Mean Squared Errors) for Additional Attributes of the Sender of an E-mail Message for the Main Effects and Interaction.

	Message Type		
Attribute	Polite/Impolite	Grammatical/Ungrammatical	P/I x G/U
Relative Power	27.40 (0.93)	116.34 (1..98)	6.80 (1.11)
Work With	87.26 (1.70)	138.06 (1..91)	143.09 (1.15)
Friendliness	239.09 (1.33)	18..54 (1.14)	124.96 (0.89)
Liking	172.36 (1.36)	29.36 (1.11)	88.42 (0.91)
Editing	17.16 (0.97)	1096.79 (1.82)	12.71 (1.25)
Reaction	40.80 (3.37)	152.13 (2.71)	48.08 (1.58)

df for all analyses (1,110) $F_{crit} = 3.93$, $p < .05$ and $F_{crit} = 6.88$, $p < .01$

Polite messages were viewed as having been written by a more friendly person as were grammatical messages. The authors of messages that were both polite and grammatical were perceived as being significantly more friendly than other messages types. (See Tables 2, 3, and 4.)

The same pattern was seen in the question about the likeability of the sender. The authors of the polite messages were seen as more likable, as were the authors of the grammatical e-mail. Furthermore, the authors of polite/grammatical messages were perceived as significantly more likable than the other combinations.

Editing refers to the extent to which the recipient thought that the message had been edited before it was sent. Polite messages, grammatical messages, and the polite/grammatical combination all were perceived as having been edited when compared to the other types.

The attribute known as "reactions" examined whether or not the recipient of the message believed that the sender of the message cared how the recipient was going to react to the message. As has been seen several times, participants saw the sender as being more concerned with the recipient when the messages were polite, grammatical, or both.

DISCUSSION

This research examined whether characteristics such as the nature of the request and the format of the request would effect the way a sender of e-mail would be perceived by the recipient of that message. Significantly consistent results were found for both the main attributes examined (power, gender and competency) as well as a set of additional attributes (relative power, willingness to work with, friendliness, likeability, degree to which the message had been edited, and reaction to the message).

The first finding that should be of interest to those doing research via electronic methods is that this study found no reliable differences on any of the variables examined between those who responded electronically and those who completed a paper-and-pencil questionnaire. Although this result is contrary to that suggested by Sproull (1986; Sproull & Kiesler, 1986), who suggested that the relative anonymity of e-mail should lead to more extreme responses, it is consistent with recent research. For example, Penrod and O'Neil (2000) found no differences between web-based and pencil-and-paper questionnaires and Schuster, Birky, and Sharkin (2000) present data that questions the assertion of more self-disclosure in electronic communication when compared to more traditional methods. Perhaps the self-disclosure difference between face-to-face versus computer-mediated communication has disappeared in the past 15 years. Or perhaps the differences can be explained in terms of the samples used—Sproull and her colleagues used employees of a Fortune 500 company while the current study was conducted with college students as participants.

Contrary to Sproull and Kiesler's (1986) results, impolite messages were seen by the participants in this study as originating with people of higher status or power. As with the method of presentation discussion, however, the nature of our sample must be kept in mind. Students use e-mail primarily for social discourse and are likely to have had little experience receiving mail from individuals with different degrees of status and power. Therefore, they could quite likely have made use of social schemas or stereotypes when responding to this question. Since impoliteness is often associated with people of high status in film and on television, the students may have replied accordingly.

On the other hand, competency was judged to be higher in both polite and grammatical messages. This is somewhat odd given the findings on status since there is some research (Lee, 1996) that suggests that status and competency are linked. However, they are in agreement with the material presented by Sherwood (1998) who maintained that the clarity of a person's communication in e-mail effects the assumptions that the recipient makes about the sender of that message. Polite and grammatical messages are likely to be seen as being clearer than impolite and ungrammatical e-mail. In addition, Garton and Wellman (1995) noted that people become less efficient when they use email. Hence, as efficiency levels drop grammatical accuracy is also apt to drop. However, more competent individuals should be able to maintain their use of correct grammar, even when using e-mail.

Messages that were impolite were also seen as being authored by males, while females were the perceived authors of the polite messages. These results are consistent with the work of Lakoff (cited in Rodino, 1997). Lakoff contends that women's language is extremely polite and traditionally women have been stereotyped as being mild mannered, while men are stereotypically less mannered and more aggressive.

As predicted, grammatical messages were seen as being authored by women, while males were perceived as authoring the ungrammatical messages. This agrees with Feingold (1993), who reported that female children perform better than male children on verbal tasks, including reading, spelling, and grammar. Therefore, when asked to make a judgement based solely on the content of an e-mail message, we are more likely to see grammatical messages as originating from a woman than a man.

In addition to the major variables examined in this study, six additional impressions of the sender were elicited from the participants. These included issues of relative power and working preferences, how friendly and likable the sender was, as well as perceptions about the care the sender took in editing the message and whether the sender cared how the recipient was going to react to the message. For all of these attributes except relative power, the polite message was viewed more positively than the impolite message. That is, people who send polite e-mail messages are seen as being more friendly and likable; recipients want to work with them; the sender is seen as having edited his or her message and is interested in the reaction of the recipient to the e-mail. The same results were seen on all attributes, even the relative power of the sender, when grammar was examined. That is, those sending grammatical messages were perceived as being friendlier and likable, recipients want to work with them, and the sender is seen as having edited his or her message and is interested in the reaction of the recipient to the e-mail. In addition, grammatical messages were seen as having originated from individuals of higher relative power or status when compared to the recipient.

The one exception to this trend was that impolite messages were seen as coming from individuals who had more relative power than the recipient. Since this result is consistent with the main findings on status, it appears that both questions were tapping into the same attitude. (Recall that one question asked if the sender has a position of power in the company while the second, the one discussed here, asked if the sender had more power than the recipient did.)

Given these main effects, the expected pattern appeared in the interactions. If polite messages are seen in a more favorable light and grammatical messages are as well, those messages that are both polite and grammatical should receive the highest ratings, and they did. The lone exception concerns relative power. Here the highest ratings were seen with messages that were grammatical but impolite.

These results indicate that in addition to being polite, senders of e-mail should be clear, i.e., use proper grammar, in what they are saying. The attributes addressed here clearly play a role in determining how people are perceived when there are few contextual cues present. And despite the dramatic increases in speed and carrying capacity, e-mail (and other means of computer mediated communication) will continue for some time in the future to provide many fewer cues as to status, gender, etc., than face-to-face communication.

Despite the strong and seemingly straightforward findings of this research, a number of potential threats to ecological validity must be acknowledged. The sample was relatively small and homogeneous and consisted of a convenience sample of college students. In addition, division into paper-and-pencil and e-mail groups was done on a self-selection basis rather than randomly (although the consistency of the results between the two groups does suggest that on the measures take here, at least, the groups did not differ). And, of course, the college students were asked to "imagine that they had received the e-mail at their place of work." Perhaps the use of these artificial stimuli in this contrived situation should be considered an interesting finding but one that needs to be explored and validated in real-life situations involving working men and women.

REFERENCES

Adkins, M., & Brashers, D. E. (1995). The power of language in computer-mediated groups. *Management Communication Quarterly*, *8*, 289-322.

Allen, B. J. (1995). Gender and computer mediated communication. *Sex Roles*, *32*, 557-563.

Burgoon, M., & Miller, G. R. (1985). An expectancy interpretation of language and persuasion. In H. Giles & R. N. St.Clair (Eds.), *Recent advances in language, communication, and social psychology* (pp. 199-229). Hillsdale, NJ: Erlbaum.

Carnevale, P. J., & Probst, T. M. (1997). Conflict on the Internet. In S. Kiesler (Ed.), *Culture of the Internet* (pp. 233-261). Hillsdale, NJ: Lawrence Erlbaum.

Chapman, G. (1995). Flamers. *The New Republic*, *212*, 13-15.

Eagly, A. H. (1987). *Social differences in behavior: A social-role interpretation.* Hillsdale, NJ: Lawrence Erlbaum.

Feingold, A. (1993). Cognitive gender differences: A developmental perspective. *Sex Roles*, *29*, 91-112.

Fisk, S. T., & Taylor, S. E. (1991). *Social cognition*. New York: McGraw-Hill.

Garton, L., & Wellman, B. (1995). Social impacts of electronic mail in organizations: A review of the research literature. *Communication Yearbook*, *18*, 434-453.

Glassberg, B. C., Kettinger, W. J., & Logan, J. E. (1996). Electronic communication: An ounce of policy is worth a pound of cure. *Business Horizons*, *39*, 74-80.

Holmes, J. (1995). *Women, Men and Politeness*. New York: Longman.

Lea, M., & Spears, R. (1991). Computer-mediated communication, deindividuation and group decision-making. *International Journal of Man-Machine Studies*, *34*, 283-301.

Lee, G. B., (1996). Addressing anonymous messages in cyberspace. *Journal of Computer Mediated-Communication* [On-line], *2*(1), Available: http://www.ascuse.org/jcmc/vol12/issue1/anon.html.

McCormick, N. B., & McCormick, J. W. (1992). Computer friends and foes: Content of undergraduates' electronic mail. *Computers in Human Behavior*, *8*, 379-405.

NUA Internet Surveys. (1998). How many online? Available: http://www.nua.net/surveys/how_many_online/world.html.

Penrod, S. D., & O'Neil, K. M. (2000). Investigating the validity of a new method of jury research: Imposing the death penalty over the world-wide web. Paper presented at the meeting of the Society for Computers in Psychology, New Orleans, Louisiana.

Rodino, M. (1997). Breaking out of binaries: Reconceptualizing gender and its relationship to language in computer-mediated communication. *Journal of Computer-Mediated Communication* [On-line], 3(3). Available: http://www.ascusc.org/jcmc/vol3/issue3/rodino.html.

Schuster, K., Birky, I., & Sharkin, B. (2000). Evaluating differences in self-disclosure based on methods of response: Internet responses vs. paper-and-pencil responses. Paper presented at the meeting of the Society for Computers in Psychology, New Orleans, Louisiana.

Sherwood, K. D. (1998). "A beginner's guide to effective e-mail." Available: http://www.webfoot.com/advice/mail.top.html.

Sproull, L. S. (1986). Using electronic mail for data collection in organizational research. *Academy of Management Journal, 29,* (1), 159-169.

Sproull, L. O., & Kiesler, S. (1986). Reducing social context cues: Electronic mail in organizational communication. *Management Science, 32,* 1492-1512.

Stahelski, A. J., & Paynton, C. F. (1995). The effects of status cues on choice of social power and influence strategies. *The Journal of Social Psychology, 135* (5), 553-556.

Weisband, S. P., & Reining, B. A. (1995). Managing user perceptions of email privacy. *Communication of the ACS, 38*(12), 40-48.

Welch, J. (1997). Electronic menaces are a flaming liability. *People Management, 3*, 14. Witmer, D., & Katzman, S. (1997). On-line smiles: Does gender make a difference in the use of graphic accents? *Journal of Computer-Mediated Communication* [On-line], *2* (4). Available: http://www.ascusc.org/jcmc/vol2/issue4/witmer1.html.

How to Submit Papers to *NA*JP

The submission of good papers is very important to any psychology journal. We hope that you will support our views about what a psychology journal should be like by submitting papers now. We will try to mail a decision on your paper within 5 weeks of receipt by us. We will publish your paper in an average of 5 months of receipt of a final accepted version.

Submit 1 paper copy of each manuscript prepared according to APA guidelines. Also include a diskette with your manuscript or (preferably) e-mail the file to NAJP@mpinet.net. Include 2 large stamped envelopes so we can send your manuscript out for review. We encourage you to have your paper read by peers before submission.

If the final paper is accepted, please send a check or purchase order for the amount specified with your final draft (billed at $25 dollars per typeset page in units of two pages).

Submission of a manuscript to the Journal is a contract to publish with the Journal if the paper is accepted. Because of our rapid editorial feedback, simultaneous submission to other journals is not allowed. We will offer brief comments on the suitability of a topic for publication in *NA*JP before a submission.

Send all correspondence to: *North American* Journal of Psychology, 240 Harbor Dr., Winter Garden, FL 34787, phone 407-877-8364.

SUBSCRIPTIONS

All subscriptions are on a calendar year basis only. If you order late in the year, the earlier issues will be sent to you. Institutional rate, $65. Individual rates, paid by check, $30. Add $7 for Canadian subscriptions. Add $8 for subscriptions outside North America. Phone orders accepted with Visa, MasterCard or American Express. Call 407-877-8364. The *North American* Journal of Psychology is now published semi-annually in June & December.

Visa ☐ AMEX ☐ MC ☐

Card # ______________________ Exp. Date ________

Name ______________________

Address ______________________

City ______________________ State ________ Zip ________

A Summary of Our Philosophy

The ***North American* Journal of Psychology** publishes scientific papers that are of general interest to psychologists and other social scientists. Papers may encompass a wide variety of topics, methods, and strategies. It is our belief that many good scientific papers are rejected for invalid reasons. Some journals reject up to 90% of their submissions because they don't have space. Furthermore, studies of the review process strongly suggest that not all of the lucky 10% are better papers than those that get rejected. In our view many of those rejected papers are meritorious with need for only minor revisions. Perhaps our philosophy can be summed up this way: we are looking for ways to strengthen good papers rather than searching for weak excuses to reject them.

North American
Journal of Psychology

Vol. 3, No. 3 2001

BEST COPY AVAILABLE

Advisory Editorial Board

ADRIAN FURNHAM
Psychology
University College London

DOUGLAS JACKSON
Psychology
University of Western Ontario

JAMES M. JONES
Psychology
University of Delaware

HEIDI KELLER
Psychology
University of Osnabrueck

MANUEL LONDON
Psychology
SUNY, Stony Brook

KAREN LYNESS
Management
Baruch College, CUNY

ROBERT PASNAK
Psychology
George Mason University

RICHARD RYAN
Psychology
University of Rochester

PHILIP ZIMBARDO
Psychology
Stanford University

Editor

LYNN E. McCUTCHEON

How to Subscribe to The *North American* Journal of Psychology

All subscriptions are on a calendar year basis only. If you order late in the year, the earlier issue will be sent to you. Institutional rate: $65. Individual rate: $30. Add $7 for Canadian subscriptions. Add $8 for subscriptions outside North America. *North American* Journal of Psychology is published in March, June and December.

Now visit our website at http://najp.8m.com

North American Journal of Psychology

December, 2001 **ISSN 1527-7143** Vol. 3, No.3

Boundaries in the Mind: Past Research and Future Directions 347
Ernest Hartmann, Robert Harrison, & Michael Zborowski

Effect of Humor on Hitchhiking: A Field Experiment 369
Nicolas Gueguen

Religion's Role in the Terroristic Attack of September 11, 2001 377
Michael E. Nielsen

Effects of Indoor Foliage Plants on Subjects' Recovery from Mental Fatigue 385
Seiji Shibata & Naoto Suzuki

Adolescent Substance Abuse: An Empirical Examination of the Covalent Security Attribution Model of User & Nonuser Groups 397
David Daniel Bogumil

Conjunctive Forms & Conditional Inference in Questions & Statements 415
Marc E. Pratarelli & Adam L. Lawson

Monitoring Children's Behaviour in a Remote Community Before and Six Years After the Availability of Broadcast TV 429
Tony Charlton, Ronald Davie, Charlie Panting, Mick Abrahams, & Lilla Yon

The Self-reported Psychological Well-being of Celebrity Worshippers 441
John Maltby, Lynn E. McCutcheon, Diane D. Ashe, & James Houran

Effects of a Health and Relationship Education Program on Drug Behaviors 453
Joseph Donnelly, Heather Ferraro, & Carolyn Eadie

Attribution and Reciprocity in International Relations: The Attribution Reciprocity Model 463
David Daniel Bogumil

Predicting Performance in 'Go' Situations: A New Use for the Cognitive Failures Questionnaire? 481
J. Craig Wallace, Steven J. Kass, & Claudia Stanny

The Development of an Attitudes Towards the Sense of Smell Questionnaire & a Comparison of Different Professions' Responses 491
G. Neil Martin, Feyishola Apena, Zinath Chaudry, Zelda Mulligan & Claire Nixon

Testosterone & Male Behavior: Research With Hamsters Does not Support the Use of Castration to Deter Human Sexual Aggression 503
Timothy O. Moore

Chandra Levy, The Missing Intern: Social, Forensic, and Evolutionary Psychology Insights 521
Russell Eisenman

Self-definition in Psychotherapy: Is it Time to Revisit Self-perception Theory? 529
Rostyslaw W. Robak

The Role of Part-set Cuing in the Recall of Chess Positions: Influence of Chunking in Memory 535
Charles J. Huffman, T. Darin Matthews, & Phillip E. Gagne

Editor's Comments

With this issue we conclude our third volume. NAJP was issued twice a year for the first two years, but this year we have completed three. I am proud of the fact that NAJP has grown in both the number of articles and page output each year. I am also pleased with the quality of the work being submitted.

Thanks again to all the reviewers, all of whom did a marvelous job of critiquing papers promptly and with great insight. Without our reviewers, this journal would simply not have the quality it has today.

This issue features an article by Ernest Hartmann and his colleagues. Dr. Hartmann is famous for his work on thin and thick boundaries. Forensic psychologist Russell Eisenman offers his insights on the Chandra Levy case, and Mike Nielsen, an acknowledged expert on the psychology of religion, weighs in with words of wisdom on the relationship between terrorism and Islam.

The November issue of *Monitor on Psychology* contained a news brief that I would like to share with you. A conference paper by Hartley and Sotto compared influential papers in psychology versus a control group of less influential papers. They found that the influential papers tended to contain shorter, easier-to-read sentences. In other words, they were easier to understand. As an editor I have encouraged clear communication and discouraged the temptation to show off one's vocabulary. I am hardly a neutral observer, but I think that NAJP authors have succeeded admirably in producing articles that are quite readable. Thanks to all the authors who have made NAJP's first three years so successful.

Lynn E. McCutcheon, editor

Boundaries in the Mind: Past Research and Future Directions

Ernest Hartmann
Tufts University

Robert Harrison
Boston University

Michael Zborowski
SUNY at Buffalo

Boundaries of one kind or another are of obvious importance in our lives. What we will discuss here is the concept of the thickness of boundaries — thin versus thick boundaries as a dimension of personality. The concept was developed as a dimension of personality, and we will review many studies relating thickness of boundaries with a variety of other personality measures, as well as other psychological measures, in normal as well as clinical populations. We will discuss a great deal of work relating boundaries to measures of sleep and dreaming. We will examine boundaries in clinical populations and in persons who have unusual mystical or paranormal experiences. We will discuss the fact that although thick vs. thin boundaries is basically a personality dimension or trait measure, there are nonetheless also intra-individual differences in boundary functioning. We all function in a more thick boundary manner at certain times and in a more thin boundary manner at other times. We will address, or at least approach, an understanding of the biology of boundaries, and we will examine some broader implications of the concept of boundaries.

The Concept of Boundaries in the Mind — Thin and Thick Boundaries

The basic underlying notion is a fairly obvious one. No matter how we think of the content of our minds — whether we think in everyday terms of thoughts, feelings, memories; in cognitive psychology terms of perceptual, semantic and memory processes (or "modules"); or perhaps in psychoanalytic terms of ego, id, superego, defenses, etc. — we are speaking of parts, or regions or processes, which in some sense can be considered separate from one another, and yet which are obviously connected. The boundaries between them are not absolute separations.

Author info: Correspondence should be sent to: Dr. Ernest Hartmann, Tufts University School of Medicine, 27 Clark St., Newton, MA 02459.

North American Journal of Psychology, 2001, Vol. 3, No. 3, 347-368.
© NAJP

The boundaries can be relatively thick or solid on the one hand, and relatively thin or permeable on the other hand.

Psychologists have discussed and explored many different aspects of boundaries. These include perceptual boundaries, boundaries related to thoughts and feelings, boundaries between states of awareness or consciousness, sleep-dream-wake boundaries, boundaries related to memory, body boundaries, interpersonal boundaries, boundaries related to sexual identity and other forms of identity, group boundaries, and boundaries in opinions and judgments. All of this has been discussed in great detail elsewhere (Hartmann, 1991).

The concept of thick versus thin boundaries as a personality measure becomes most clear if we examine the many kinds of boundaries, as in Table 1, and consider extreme examples for clarity. A person who has very thick boundaries in all senses would be someone with a sharp sense of focus, who can easily concentrate on one thing while ignoring others. This person does not experience synesthesia, keeps thoughts and feelings entirely separate ("I don't let my feelings get in the way of my thinking"), and is absolutely clear about when s/he is awake, or asleep or dreaming, experiencing no in between states. This person has a clear sense of the separation of past, present, and future ("that was then, this is now"), has a very definite sense of space around him/herself ("this is my space, this is yours"), and has a very clear, delineated sense of sexual identity ("I am a man, you are a woman, vive la différence"), group identity ("this is my group, we do such and such; other groups are totally different) and tends to see the world in terms of black versus white, us versus them, good versus evil.

A person at the other extreme, a person with thin boundaries in all senses, may experience synesthesia, will tend to let a lot of sensory material in at once, and may have difficulty focusing on one part of the input. This person will tend to be aware of thoughts and feelings together ("I can't imagine a thought without a feeling"), and will often experience states of being half-awake and half-asleep. Or this person will become deeply immersed in daydreaming or in reverie, so that sometimes the boundary between real life and fantasy may be unclear. There will be less sense of clear body boundary and personal space. This person may be very aware of the past, and have it blend with the present ("I am grown-up, but in a lot of ways I'm still a child"). Similarly, this person will accept mixtures in sexual identity ("I am a man, but there's a lot of feminine in me too"). He or she will not feel solidly a member of one group, but rather be an individual taking part transiently in many groups, or perhaps a "citizen of the world." In judgments or opinions about the world, this person will tend to think in terms of shades of gray, rather than black and white ("it all depends, s/he's good in some ways and bad in others," "it's different at different times," and so on).

These, ofcourse, are extremes. Most of us are somewhere in between, a mixture of thin and thick boundaries, but this gives a flavor of the personality dimension running from very thick to very thin, which has recently been quantified using the Boundary Questionnaire (see below).

Precursors of the Boundary Concept

Nothing under the sun is entirely new. The concept of thin and thick boundaries is related in some way to a number of previous dimensions and dichotomies. For instance, William James (1907) divided people into "tough minded empiricists," and "tender minded rationalists." Kurt Lewin, in the 1930s, diagramed the mind as a number of regions acting on one another, separated by divisions of various thickness (Lewin, 1936). Freud discussed boundaries only a few times, especially when he speaks of the stimulus barrier or "reitzschutz" — a protective shield against stimulation. He referred to the entire ego as initially a body-ego derived from the body surface (Freud, 1923). Many of Freud's followers did explore boundaries in more detail (see for instance Federn,, 1952). There is an entire literature on "ego boundaries" which definitely form part of what we are speaking of here. In the psychoanalytic literature, solid ego boundaries are considered a kind of ideal, and the emphasis is on defects and weaknesses in ego boundaries which lead to psychosis or other pathological conditions (this is quite different from the view of thin and thick boundaries as a value-free personality dimension, which we develop below). A French psychoanalyst, Anzieu, has worked clinically with the concept of the "ego skin" (moi pau) as an "envelope for the ego," (Anzieu, 1987). He is obviously speaking of boundaries too.

Clinical psychoanalysts have generally made no attempts to quantify these boundary measures. Such attempts have, however, been made by such as Blatt, and Ritzler (1974), using the Rorschach. Peter Landis has studied ego boundaries in detail and developed some ingenious tests for ego and interpersonal boundaries (Landis, 1970). All of these measures can be related to thin versus thick boundaries. Fisher and Cleveland (1968) have worked extensively with two measures, "Barrier," and "Penetration," based on the Rorschach test. Theoretically, "Barrier" ought to be closely related to thick boundaries and "Penetration" to thin boundaries. However, empirically, this is not the case. The "Barrier" and "penetration" measures turn out not to be opposites (Fisher and Cleveland, 1968), and further, neither seems closely related to thick and thin boundaries (Fisher, 1992, unpublished manuscript).

Rokeach (1960), in his work on the "open and closed mind," was clearly dealing with an aspect of boundaries, as were Adorno and his colleagues in their classical work on the "authoritarian personality" (Adorno, Frenkel-Brunswik,, Levinson,. & Sanford, 1950). The "closed mind" and the "authoritarian personality" definitely describe aspects of people with very thick boundaries.

Finally, thick and thin boundaries may be relevant to different styles of organizing mental contents. In different ways, Mednick (1962), Spence (1964) and Broadbent (1971) distinguish between a conscious, logical, hierarchical style of conceptual organization, on the one hand, and a preconscious, connotative, parallel processing style of conceptual ganization on the other. Each style may serve important defensive as well as adaptive purposes. By being neat, explicit, and well organized, people with thick boundaries can reduce the chances of different concepts becoming confused with each other; perhaps at the cost of not seeing novel connections between them (Mednick, 1962). Thick boundaries can be used defensively to avoid seeing connections between related ideas. While thin boundaries between concepts permit novel and sometimes creative associations between normally unrelated ideas, thin boundaries may be implicated in confused and autistic thinking. In this regard, a cognitive style, category width, (Gardner, Holzman, Klein, Linton,, and Spence, 1959) has to do with the number of diverse objects a person can tolerate as belonging to the same category or group. To consider two different things as belonging to the same group, the conceptual boundaries between them must be relaxed. Thus, we believe that thin and thick boundaries represent an important and pervasive personality dimension.

The Boundary Questionnaire (BQ)

The BQ is a 138-item questionnaire including items about many different aspects of boundaries (Hartmann, 1989, 1991; Hartmann,, Harrison, Bevis, Hurwitz, Holevas, Dawani, 1987). Table 1 gives some illustrative items, divided into 12 categories. The *Boundary Questionnaire* has now been "taken" by over 2000 persons in our own studies, and several thousand more in a number of other investigations. The response format for each question runs from '0' (= not at all) to '4' (= very much so). Approximately two thirds of the items are phrased so that full endorsement (very much so) indicates a 'thin' boundary, and the remaining items are phrased so that 'very much so' indicates a thick boundary. To score the test, the answer-values of the thick 'items' are reversed, and all of the scaled answers are added to produce a Sumbound score. In a sample of 866 participants, gathered from various sources (specific thin and thick-boundary groups, college students) the correla-

TABLE 1 Boundary Questionnaire: Examples of Items

- Category 1: Sleep/Dream/Waking
  - 1. When I awake in the morning, I am not sure whether I am really awake for a few minutes.
- Category 2: Unusual Experiences
  - 61. At times I have felt as if I were coming apart.
- Category 3: Thoughts/Feelings/Moods
  - 15. Sometimes I don't know whether I am thinking or feeling.
- Category 4: Childhood/Adolescence/Adult
  - 4. I am very close to my childhood feelings.
- Category 5: Interpersonal
  - 103. I am a very open person.
- Category 6: Sensitivity
  - 42. I am unusually sensitive to loud noises and bright lights.
- Category 7: Neat/Exact/Precise
  - 19. I keep my desk and work table neat and well organized.
- Category 8: Edges/Lines/Clothing
  - 32. I like heavy, solid clothing.
- Category 9: Opinions regarding Children, etc.
  - 56. I think a good teacher must remain in part a child.
- Category 10: Organizations
  - 58. A good relationship is one in which everything is clearly defined and spelled out.
- Category 11: Peoples/Nations/Groups
  - 11. People of different nations are basically very much alike.
- Category 12: Beauty/Truth
  - 36. Either you are telling the truth or you are lying; that's all there is to it.

correlations of Sumbound with all of the items were positive. The alpha reliability for the test is .925. All 138 items load positively on the first principal component, and the Armor theta reliability (Armor, 1973-4) is 0.927. Since 'agreement-set,' the tendency to agree, was controlled by reversing the scoring direction of one third of the questions, the uniformly positive loadings attest to the idea that there is one (over-arching) principle underlying subjects' responses to all 138 questions. The BQ has good test-retest reliability over six months (r's of about .77 in two samples and Kunzendorf & Mauerer 1988-89, Funkhauser, Würmle, Comu, & Bahro 2001).

An exploratory factor-analysis was done on the correlations among the 138 questions, using principal-components factor-extraction. Using Cattell's (1946) 'scree' test, and subsequent interpretability as criteria, thirteen factors, accounting for 37.3% of the variance, were preserved for rotation using the Normal Varimax criterion. The 13th eigenvalue was

1.65. Items loading 0.25 or above on a given factor were regarded as belonging to it. The first 12 rotated factors were easily interpreted; the 13th was un-interpretable, its 8 items accounting for less than 1% of the total variance. The content of each factor appears in Table 2 (Harrison et al., 1989).

TABLE 2 Summary of the Factor Analysis (from Harrison et al., 1989)

Factor I-13, Primary Process Thinking:
The 51 items in this factor (all keyed in the 'thin' boundary direction) describe a person who has many experiences of merging; of fluctuating identity; whose imagery is so vivid it is hard to distinguish from reality; who experiences synaesthesia; the merging of objects with self and with each other. 49 of the items are keyed 'True'. Theta reliability (see Armor 1973-4) = .92.

Factor II-13, Preference for Explicit Boundaries*:
The 37 items on this factor (36 keyed in the 'thick' boundary direction) express a preference for clear borders whether it is in nation, cities, houses, pictures, stories, or relationships. A secondary emphasis is on neatness. 34 of the items are keyed 'True'. Theta = .87.

Factor III-13, Identification with Children:
The 19 items in this factor (18 keyed 'thin') describe a person who feels, in part, like a child; identifies with children and enjoys them. All items are keyed 'True'. Theta = .75.

Factor IV-13, Fragility:
The 13 items in this factor (12 keyed 'thin') express sensitivity to hurt, a difficult and complicated childhood and adolescence, fears of falling apart and fears of being overwhelmed by interpersonal involvement. 12 keyed 'True'. Theta = .75.

Factor V-13, Clairvoyance:
The 16 items on this factor (14 keyed 'thin') include beliefs in one's clairvoyant powers including knowing others' unexpressed thoughts and feelings, having premonitory dreams, and experiencing very vivid memories and imagery. These items also suggest a strong sense of self-identity from childhood through old age. 15 of the items are keyed 'True'. Theta = .70.

Factor VI-13, Open-ness:
The 11 items on this factor (all keyed 'thin') describe a person who believes in being open to the world, trusting others, and disclosing personal experience. 10 of the 11 items are keyed 'True'. Theta = .70.

To determine the stability of the factor solution, we re-factored the BQ for 364 college students in the sample only, and found an almost-identical factor-structure. Factor-loadings for this sub-sample were within 0.02 of those found for the total group (Harrison et al, unpublished). Even more recently, a new factor analysis on 500 students has replicated the original results with an almost identical factor structure (Zborowski, 2001, Personal Communication).

Boundary scores can be obtained for each of the individual categories and factors. However, the most used measure has been the overall boundary score called Sumbound, in which high numbers signify thinness. In the first 1000 subjects studied, the range has been 120-454, the mean value 271 ± 50.

Who has Thick or Thin Boundaries

A number of interesting findings have emerged as to who may be characterized by thick or thin boundaries. First, even though the items were very carefully written to have no gender bias, women consistently score significantly "thinner" (one half of a standard deviation) than men, and there is also a slight age effect: older subjects score slightly thicker than younger subjects (Harrison et. al.. 1987 Personal Communication, Hartmann, 1991). However, no long-term studies have been done as yet to determine how boundaries develop and change over the years within a single person.

Significantly thinner boundaries compared to control groups have been found in art students (Beal, 1989, Hartmann, 1991), music students, and mixed groups of creative persons (Beal, 1989), frequent dream recallers (Hartmann, 1991, Hartmann Elkin, & Garg 1991), adults with nightmares (Hartmann, 1991, Levin, Galin, & Zywiak 1991; Galvin, 1993), adolescents with nightmares (Cowen and Levin, 1995), "lucid dreamers" (Galvin, 1993), male as well as female fashion models (Ryan 2000), persons with unusual mystical experiences (Krippner., Wickramasekera, Wickramasekera, & Winstead, 1998), and persons with a diagnosis of Borderline Personality Disorder, Schizoid Personality Disorder or Schizotypal Personality Disorder (Hartmann, 1991). Interestingly, although art students have much thinner boundaries than average, this is not true of established artists, who have boundary scores in the normal range (Beal, 1989).

Groups that score significantly "thicker" than average on the BQ include naval officers, salespersons, lawyers, patients with a diagnosis of Obsessive-compulsive Personality Disorder, persons suffering from "Alexythymia" (Hartmann, 1991), and patients (from two different sleep disorders centers) with a diagnosis of Sleep Apnea (Hartmann, 1992).

The Relationship of the BQ to Other Personality Measures

When the BQ was first used in 1985, it appeared to be a new dimension of personality, not clearly related to any of the then standard personality measures. Thus, there are only low and non-significant correlations between BQ and Eysenck's personality dimensions, although one study found some relationship between thin boundaries and Neuroticism in a small group (Sand and Levin, 1996). There were also no clear relationships to Cloninger's three dimensions of personality.

The BQ did show some relationships with MMPI scales (Hartmann, 1991). In 299 subjects, relationships found were very consistent with what we had predicted on the basis of the definition of boundaries. Sumbound correlated positively ($r = .32$) with the F ("atypical response") scale, and this appeared to be a valid relationship. Subjects scoring thinner on the Boundary Questionnaire did frequently report and discuss the unusual experiences described on the F scale, for instance, "I have a nightmare every few days." Sumbound showed a negative relationship ($r = -.37$) with the K scale ("defensiveness,"), which can be considered an aspect of thick boundaries. Sumbound correlated positively ($r =.41$) with Pa (paranoia), which is not surprising, since it is accepted that Pa in normal groups measures a kind of sensitivity rather than blatant paranoia. Finally, Sumbound correlated positively ($r = 0.40$) with the Mf scale in males — consistent with the view that thin boundaries involves the ability for males to be interpersonally sensitive, and to see feminine elements in themselves. Although these were highly significant correlations, all $p < .001$, the modest size of the correlation suggests that the BQ is obviously measuring something different than these individual MMPI scales.

Significant positive correlations have been reported between Sumbound on the BQ and several measures of hypnotizability and suggestibility (Barrett, 1989, Rader, Kunzendorf, and Carrabino 1996), as well as measures of creativity (Levin, Galen, & Zywisk 1991). An especially strong correlation ($r = .67$) has been found between Sumbound and Tellegen's *Absorption Scale* (Barrett 1989). Again, these relationships were as predicted from our description of thin boundaries, above.

On the *Rorschach* test, subjects with thinner boundaries were found to have significantly higher boundary disturbance scores, and also significantly lower form quality scores (Levin, Gilmartin, & Lamontonaro 1998-1999). Recent studies have established a relationship between thin boundaries and a number of other measures relating to personality, including certain forms of anxiety. An especially strong relationship is found between Sumbound and Insecure Attachment, measured on the *Bell Object Relations and Reality Testing Inventory* (Bell , Billington & Becker 1986). (Hartmann and Zborowski, 2001). Thin boundaries are also positively related to measures of connection-seeking, at least in women (Bevis, 1986). And there is a high correlation ($r = .51$) between thin boundaries and rated openness in an interview study (Zborowski, Hartmann, & Newsom 2001, Manuscript submitted for publication).

There have been two separate investigations relating the *Boundary Questionnaire* to the *Myers-Briggs Inventory*. In both studies the most striking finding was a positive correlation (r between .4 &.5) between Sumbound and "Intuition," and a somewhat smaller correlation with "Feeling" (Erhman and Oxford, 1995, Barbuto & Plummer 1998, 2000).

A few preliminary studies suggested that the BQ was unrelated to Norman's basic Five-Factor structure of personality. However the Five Factor Model has evolved, and the more recent model championed by Costa and McRae (1992), includes, as one of the five dimensions, "Openness to Experience." McRae (1994) has recently reported a very high correlation ($r = .73$) between thinness of boundaries on the BQ (Sumbound) and Openness to Experience. We have attempted to further examine this surprisingly high correlation. Indeed the Boundary Questionnaire includes at least two items "I am a very open person" and "I am a very sensitive person" which plainly relate to items in "Openness to Experience." And in fact, factor VI of the BQ was named "open-ness" long before the relationship of the BQ to "Openness to Experience" was known. A detailed examination of the items in the "Openness to Experience" scale is also revealing. The items involve several aspects of boundaries, but emphasize the desirable or positive aspects of thin boundaries. For instance, "I have a lot of intellectual curiosity," "I often enjoy playing with theories or abstract ideas," and (scored negatively) "I have little interest in speculating on the nature of the universe or the human condition." Openness to Experience does *not* include any of the less attractive aspects of thin boundaries, such as feeling overwhelmed by input, vulnerability, becoming over-involved in a maladaptive way, etc. Thin Boundaries and Openness to Experience are obviously closely related, but in our opinion thick versus thin boundaries represents a broader and perhaps more useful measure since it is neutral or value-free and covers both adaptive and maladaptive features.

In this connection it is interesting that BQ shows close to zero correlation with the *Marlowe-Crowne Social Desirability* scale (Earle 1992). Overall, neither thin nor thick boundaries are considered more desirable than their opposite. However, a careful examination of the answers and a series of interviews has convinced us that by and large people consider their own type of boundary structure as most desirable. Thus, people with very thick boundaries tend to use terms for others with thick boundaries such as "solid," "reliable," "lots of perseverance," etc., while they characterize people with thin boundaries as "flaky," "far out," "unreliable." People who themselves score very thin on the BQ speak of those with thick boundaries as "dull," "rigid," "unimaginative," while they think of those with thin boundaries as "exciting," "creative," "innovative."

Thin Boundaries and Unusual Sensitivities

There are a number of suggestive studies indicating that people with thin boundaries may be not only creative and open, but may have a series of other interesting and so far poorly understood characteristics. For instance, there appears to be a relationship between thin boundaries and multiple chemical sensitivities (Jawer, 2001). There is also a correlation between thin boundaries and a belief in or tendency to experience paranormal phenomena. Factor V of the BQ appears to pick up this aspect of thin boundaries and has been labeled "clairvoyance." Groups of people who characterize themselves as shamans or psychics score thin on the BQ (Krippner, Wickramasekera, Wickramasekera, & Winstead, 1998). Thalbourne and his collaborators, in their studies of persons who experience paranormal phenomena, have devised a "Transliminality scale" to measure these traits (Lange, Thalbourne, Houran, & Storm 2000; Thalbourne, 1991). Preliminary analysis suggests a high correlation ($r = .68$) between thin boundaries and the Transliminality Scale.

These relationships may be worth exploring further, since two very different hypotheses may explain them. The most parsimonious view would be that all "paranormal" phenomena are imaginary, and that people with thin boundaries simply have better or looser imaginations, are more suggestible, or are more sensitive with a tendency to elaborate creatively on their sensitivities. On the other hand, we could consider the possibility that phenomena such as telepathy, now considered paranormal could be related to transmission of information using perhaps portions of the electromagnetic spectrum which we are not usually able to detect. Under unusual circumstances our ability to detect such information could be altered slightly, and quite possibly there might be inter-individual differences in the ability to detect information of this kind. If so, it is possible that persons with thin boundaries who are sensitive in so many other ways, may also be sensitive to detecting such portions of the spectrum.

Along these lines, psychologist Daryl Bem, who was until recently a skeptic about any kind of paranormal experiences, carefully reviewed all quantitative studies of these phenomena and came to the conclusion that there is indeed something worth studying. Bem and Honorton (1994) in a paper called "Does PSI Exist?" interpret the overall statistical results as suggesting that there is a small positive effect, especially in studies of telepathy and clairvoyance, which they can not attribute to chance, poor study design, etc. They also comment that the results are most positive when the subjects used are artists or creative people (in other words, probably people with thin boundaries). Certainly nothing is proven, but we suggest that anyone considering a careful quantitative study of "psi" or paranormal phenomena should consider using people with thin boundaries as subjects.

Clinical Boundaries — Boundaries in Psychotherapy

Although psychotherapists have been making use of concepts such as ego-boundaries and interpersonal boundaries for many years, there has not been much quantitative study or much attempt to relate therapy to the BQ. One small survey suggests that people with high Sumbound scores (thin boundaries) are much more likely to value psychotherapy, to have been involved in psychotherapy, and to have benefited from psychotherapy (Hartmann, 1996). One paper discussed thick and thin boundaries in relation to "boundary violations" in therapy. The conclusion is that victims of sexual boundary violations tended to be persons (usually women) with thin boundaries. Interestingly, it appeared that the violators of boundaries (therapists who had sex with their patients) could be divided into two distinct groups, one group with unusually thin boundaries, and another group with very thick boundaries. The first group consisted of therapists who were themselves somewhat confused and unclear about rules and limits in psychotherapy and elsewhere. The second group consisted of therapists who were not confused, but knew exactly what they were doing. They were solid, "thick-boundary" people who lacked empathy, and were not able to appreciate the pain and suffering that might be experienced by the patient involved in the boundary violation (Hartmann, 1997).

We have repeatedly found that an explicit discussion of boundaries can be very useful in psychodynamic therapy. This is especially the case in patients who have relatively thin boundaries, or at least thin boundaries in some areas, and are painfully aware of this when it is called to their attention. It is often useful for them to think of their problems in terms of boundaries being thin in certain areas, and perhaps needing thickening. Further, they appreciate the fact that they are in this sense similar to art students and other creative persons. The same patients often react poorly to being labeled with diagnoses such as "Borderline Personality Disorder," or "Schizotypal Personality Disorder" which have a great many negative connotations for the patient and sometimes the therapist.

Boundaries and Dreams

There have been a large number of studies relating boundaries to various aspects of dreams. First of all, there is a clear and significant positive relationship between thinness of boundaries and amount of dream recall in various normal and clinical samples (Hartmann, 1991, Hartmann Elkin, & Garg 1991; Hartmann, Rosen, & Rand 1998;

Kunzendorf, Hartmann, Cohen, & Cutler 1997; Levin et al 1998-99; Schredl, Kleinferchner, & Gell, 1996; Schredl, Schafer, & Hoffman 1999;). This relationship holds in an adolescent population as well (Cowen and Levin, 1995). In most studies, dream recall was measured by a single question, but a study involving dream diaries obtained similar results (Schredl et.al., 1996). This relationship has been replicated a number of times, and holds true even when all items on the BQ relating to dreams, daydreams, sleep, and waking are removed (Hartmann, 1991). In the past it has been strikingly difficult to obtain consistent relationships of dream recall with any personality measures. In fact, the positive relationship with thinness on the BQ is at present the only well-replicated correlate of dream recall. Although the results are highly significant, the magnitude of the *r*'s (usually 0.2 to 0.5) indicates that, obviously, much variance is left to be explained. Part of this variance is certainly "noise in the system" — the fact that simply asking people a single question about how often they recall a dream is a very rough measure. Many people — probably the majority — have thought little about their dreams or amount of dream recall, and their answer to the question is only a wild guess. In fact, one study showed that the correlation between dream recall and thinness on the BQ is much higher in persons who are especially interested in their dreams — members of the Association for the Study of Dreams. The correlation between BQ score and dream recall was .58 in this subgroup ($N = 42$), whereas it was .40 in the larger, unselected group of 600 persons (Hartmann, 1991).

Dream recall shows a significant positive correlation not only with the overall BQ score SumBound, but with each of the twelve categories of boundaries. Also, a group of frequent dream recallers (seven dreams or more per week) scores significantly higher on SumBound and on each of the twelve categories of the BQ than a group of persons who say they do not recall dreams (Hartmann, 1991; Hartmann, et al, 1991). Further, the BQ is related to aspects of the content of dreams: dreams of subjects with thin boundaries are scored as more vivid, more dreamlike, more emotional, and having more interactions between characters (Hartmann et al, 1991; Hartmann et al. 1998). Similar findings on dream content are reported by Schredl et al. 1996). In another study, Sumbound was positively correlated with the emotional intensity, bizarreness, and "morbid content" of dreams (Zborowski, McNamara, Hartmann, Murphy & Mattle 1998).

There is also a relationship between boundaries and the CI score (a measure of the power of imagery in the dream in which there has been considerable recent interest. CI refers to a "Contextualizing Image." For instance, a person who has recently experienced trauma of any kind often has dreams such as "I was overwhelmed by a tidal wave." The powerful tidal wave imagery is thought to contextualize or picture the underlying emotion of terror or vulnerability (Hartmann, 1996, 1998). CI scores have good inter-rater reliability, and have been found to be significantly higher in dreams than in daydreams (Hartmann, Rosen & Grace 1998), and higher in the dreams of people who have experienced trauma or abuse than those who have not (Hartmann, Zborowki, Rosen & Grace 2001). Concerning boundaries, it was found that a group of students who scored thin on the BQ had significantly higher CI scores in their written "most recent dreams" than students who scored thick (Hartmann, Zborowski, McNamara, Rosen & Grace, 1999).

Galvin (1993) found that lucid dreamers — persons who frequently realize they are dreaming during their dreams — have thin boundaries with scores in about the same range as nightmare sufferers. However, the two groups differ on factor 10 of the BQ: the nightmare group is more "over-involved."

Aside from the relationship with dreaming, there are also some intriguing preliminary findings suggesting relationships between boundaries and other aspects of sleep and sleep disorders. Overall, there is a small but significant correlation between thinness of boundaries and length of sleep (Hartmann, 1991). Indeed, persons classically described as "short sleepers" — people who get along on 6 hours or less of sleep per night, do not catch up, and do not have any complaints of insomnia — appear to have many characteristics of people with thick boundaries, though the BQ was not available at the time of the sleep studies (Hartmann, Baekland, & Zwilling 1972).

Among people with sleep disorders, we have mentioned that those who have frequent nightmares have unusually thin boundaries (Hartmann, 1984, 1991). This relationship has been studied and confirmed (Zborowski et al, 1998). However, persons who have a very different condition, night terrors — early night frightening awakenings without dreams — as well as those with sleepwalking (sometimes associated with night terrors) tend to have thick boundaries. Patients suffering from bruxism (tooth-grinding) tend to have thicker boundaries than average, as do patients with sleep apnea, as mentioned (Hartmann, 1991, 1992).

Boundary Changes Within an Individual: the Focused-Waking-to-Dreaming Continuum

We have chiefly discussed thin vs. thick boundaries as a personality trait, and in fact most studies have focused on boundaries in this way. However, it is interesting to note that thickness of boundaries may also vary within an individual at different times or in different states. We have studied this variation, especially as related to the *focused waking to dreaming* continuum as follows.

In the studies above and the usual conceptualization, dreaming is treated as though it were totally separate from waking. However, a great deal of data suggests that it is useful to think of a continuum running from focused waking on one end to dreaming at the other end. This is discussed in detail elsewhere (Hartmann, 1998). At the "focused waking" end of the continuum, the mind deals with words, mathematical symbols, logical relationships, tending to use sequential processing (A—>B—>C—>D), making connections in a serial fashion. At the dreaming end, the mind is hyper-connective, using loose associations, merging, condensation and so on, making connections in a parallel or auto-associative fashion. The continuum has a clear relationship with boundaries in the senses we have discussed above. When we are wide awake engaged in focused waking activities, such as doing an arithmetic problem, or chasing down a fly ball in the outfield, we are functioning in a thick boundary manner. We try to "think clearly" and make clear separations. We carefully separate "task" from "non-task," and one task from another. When we move towards daydreaming and eventually dreaming, categories merge, and we combine and relate things in all sorts of unusually ways. Freud's first and most important mechanism of the "dreamwork," known as *condensation*, is exactly this — a combining of separate items, a blurring of boundaries. Thus, we all have thinner boundaries when we are dreaming. In this sense, the continuum from thick boundaries to thin boundaries, can be considered not only as a personality trait, but also a within-person state measure. A number of studies suggest that content from waking, especially at the reverie or daydreaming portions of waking, are not entirely differentiable from dreaming. One study demonstrated the relationship with the BQ. It showed that overall students' dreams were scored as more "bizarre" and "dreamlike" on standard scales than students' daydreams. However, students who scored thin on the BQ showed a shift to the right relative to thicker students, so that the daydreams of students who scored thin were scored just as "bizarre" and as "dreamlike" as the dreams (night dreams) of students who scored thick (Kunzendorf et al, 1997).

Another study showed that in general subjects had more difficulty distinguishing visual pulses from auditory pulses under hypnosis than while wide awake. This difference was stronger in subjects scoring "thin" on the boundary questionnaire (Kunzendorf and Maurer, 1988-89).

Boundaries and the Brain

If thin versus thick boundaries represents a clear-cut personality dimension and also an aspect of mental state functioning, one would predict that thick versus thin boundary functioning should be detectable on the biological level, in terms of brain function and activity. In simple terms, one might suggest that thin boundaries, relative to thick boundaries, might be associated with more hyper-connectivity, perhaps a more rapid spread of activation in the forebrain, or more "in-between states." It has not been easy to study boundaries biologically, but a few beginnings have been made. For instance, in examining the polysomnograms (all-night sleep records) of persons with very thick or very thin boundaries, we were struck by the fact that people with thick boundaries appeared to have more clear-cut states of waking, NREM sleep, and REM sleep. The records of those with thin boundaries showed more in-between states, or difficult-to-define states. Several researchers have had this impression, but it has not yet been validated in a completely blind fashion with agreed-upon definitions of in-between states. Along similar lines, a study by Watson (1985) investigated phasic integrating potentials (PIPs), which are sharp spikes recorded in humans and animals, occurring chiefly during REM sleep—and in fact considered an index of the basic neurophysiology of REM sleep—but occasionally at other times as well. Watson found a strong, positive correlation (r = .52) between the number of PIPs outside of REM sleep and thinness of boundaries. In other words, people with thinner boundaries in a psychological sense, also had thinner boundaries between REM and NREM sleep: the brain activity characteristic of REM sleep often "escaped" into NREM sleep. More work along these lines is definitely needed.

Another small study investigated changes in skin temperature induced by imagining warm or cold scenes (sitting by a fire or holding an ice cube). Subjects who scored thin on the BQ showed a greater actual change in skin temperature in response to these conditions (Hartmann 1991). Similarly, in a group of 78 students, those who scored thin on the BQ showed more and longer-lasting autonomic arousal (measured by skin conductance) to an arousal-producing stimulus than other students (Levin and Fireman, 1993).

There has been one study in which a number of neuro-psychological tests were administered to persons scoring either very thin or very thick on the BQ. Such tests are generally used to detect brain damage, so of course large differences cannot be expected in these normal subjects. Yet significant differences were found in some interesting areas. Subjects scoring thick showed more evidence of perseveration (continuing in a task or strategy when asked to change) and a more systematic approach to constructing figures. Thin boundary subjects were significantly better at changing strategies or adopting new strategies on a number of tests (Garg and Hartmann, 1993). Such differences are suggestive of differences in functioning of the frontal cortex. To the best of our knowledge there have as yet been no studies using imaging techniques such as PET or fMRI, which might be expected to show such differences more clearly.

There are also definite hints of a neuro-pharmacology and neurochemistry of boundaries, though with no controlled studies as yet. Based on several hundred interviews with research subjects and patients taking a number of psychoactive medications, it is clear that certain medications produce a temporary thickening of boundaries. These include the stimulants and some anti-depressants. Persons with thin boundaries taking such medications repeatedly describe the effects as: "I feel less pulled-apart;" "I feel focused;" "I feel ready to work on things;" "I do less daydreaming;" "I'm more concentrated;" "I feel tougher;" "I can pull my thoughts together better;" "instead of being pulled in ten different directions, I can move in one straight line." These "thickening" effects appear more prominent in people who have thin boundaries, regardless of whether they are normal subjects in studies, or patients with a variety of diagnoses.

On the other hand, the effects of psychedelic drugs, especially LSD, can be interpreted as a temporary thinning of boundaries, with effects that include synesthesia, vivid changing imagery, dreamlike states, many thoughts or feelings simultaneously. One woman who scored very thin on the BQ said of an LSD experience, "I can see why some people might like or need this sort of loosening or merging, but it's not for me. I'm too much like that anyway, without drugs."

These clinical or anecdotal results suggest the probable importance of the biogenic amines in boundaries. They suggest especially that increased norepinephrine (and perhaps dopamine) in the cortex can produce a thickening of boundaries, whereas a lack of norepinephrine can produce thinner boundaries. This is consistent with the fact that the lack of norepinephrine in the cortex is one of the clearest characteristics of REM sleep, when most dreaming occurs. All of this can be considered only a small beginning to an investigation of the biology of boundaries.

Overall, thick vs. thin boundaries appears to be a robust personality measure that can be considered an important dimension of personality. This measure is related to, but may in some ways be broader and more useful than, the personality dimension "Openness to Experience." We have discussed many correlates of thick and thin boundaries, and have also emphasized the relationship of thick and thin boundaries to the continuum running from focused waking thought to dreaming. In this sense, people with thick boundaries spend more time and find themselves more comfortable at the left-hand end of our continuum involved in focused waking. They can be considered "thought people," whereas the people with thinner boundaries are more comfortable at the other end of the continuum and can be thought of as "dream people," although these terms are obviously an over-simplification.

Broader Implications of Boundaries for Society and Humanity

Beyond the implications for personality, behavior, etc. mentioned above, we would like to discuss here some broader speculations about boundaries in the mind. These are very difficult to investigate in controlled studies, and are offered principally as "food for thought."

Boundaries have a great deal to do with not only the organization of our thinking, but of our societies. For instance, if we think about law, we can see a very thick boundary position: "this is legal, that is illegal;" "this man is either innocent or guilty, that's all there is to it;" as opposed to a more thin boundary position involving shades of gray, mitigating circumstances, and partial responsibility.

This difference can be seen not only in law, but also more generally in ethics. The thick position is that there is clear-cut good and clear-cut evil, while the thinner view involves shades of gray and "situation ethics."

Quite likely aesthetics is influenced by boundaries in a somewhat similar way. The thick boundary position would be "this is beautiful, this is ugly; there are rules that tell you exactly which is which; that's all there is to it." The thin boundary position might be more like, "it's hard to tell, it all depends," or "I can see something beautiful in this even though it looks ugly overall," "it's hard to judge," or perhaps that the artist can transform ugliness into beauty. For instance, Rodin states, "what is commonly called ugliness in nature can in art become full of great beauty" (Gsell, 1971).

Other philosophical categories can likewise be seen differently from a thick or thin boundary position. Even basic building blocks of metaphysics such as "real" vs. "unreal" or "true" vs. "false" may be seen in absolute terms, or in a more relativistic fashion. In an overall sense, thick boundaries can be visualized as black vs. white, while thin boundaries can be thought of as various shades of gray.

In general, we suggest that thick boundaries are more helpful at times of war, threat, or danger. In fact societies or groups that feel threatened or in danger, for whatever reason, tend to develop thicker boundaries.* Thin boundaries are more helpful, or perhaps simply more possible, when there is less danger and we can "let go." Along political lines, it is possible to think of war and peace in very thick boundary terms. A "thick boundary peace" consists of establishing precisely the interests, the duties, the territory of the two sides, and making appropriately clear agreements: "this is mine, this is yours; we'll set up all the walls, regulations, etc. that may be necessary." A "thin boundary peace" would be more along the lines of joint occupations of areas, joint assemblies, flexibility, a gradual realization that "we" are not really so different from "them," and eventually a situation where it doesn't matter much to whom a particular bit of territory belongs. This latter position is considered far-out and idealistic by much of the world, but it has occurred within the United States or within Switzerland, where there are continuing disagreements between states (or cantons) as to who owns exactly what territory, what river rights, etc., but it is not anything to make war about. It is no longer a matter of life or death. This has also occurred recently in Western Europe, where countries which not long ago were bitter enemies, now allow citizens to cross borders without even requiring papers.

In an adaptive sense, there are advantages to thin as well as thick boundaries. For the survival of an individual, it is certainly important to be able to concentrate and engage in focused waking activity, but also at times to be able to daydream or dream. For a corporation, it may be useful to have a creative genius/inventor with an initial idea, or several at different times, but then to have more thick boundary types to develop, organize, market, etc. Likewise for society as a whole, it is perhaps optimal to contain and value individuals with thin as well as individuals with thick boundaries.

REFERENCES

Adorno, T., Frenkel-Brunswik, E., Levinson, D. & Sanford, R. (1950). *The authoritarian personality.* New York: Harper.

Anzieu, D. (1987). *Le Moi peau.* Paris: Monod Press.

Armor, D. (1973-4). Theta reliability and factor scaling. In H. Costner, (Ed.), *Sociological Methodology*, San Francisco: Jossey-Bass.

Barbuto, J., & Plummer, B. (1998). Mental boundaries as a new dimension of personality: a comparison of Hartmann's boundaries in the mind and Jung's psychological types. *Journal of Social Behavior and Personality*, *13*, 421-436.

Barbuto, J. & Plummer, B. (2000). Mental boundaries and Jung's psychological types: a profile analysis. *Journal of Psychological Type*, *54*, 17-21.

Barrett, D. (1989). The relationship of thin vs. thick boundaries to hypnotic susceptibility. Paper presented at the meetings of the Eastern Psychological Association. Boston, MA. April, 1989.

Beal, S. (1989). The boundary characteristics of artists. (Unpublished doctoral dissertation, Boston University).

Bell, M., Billington, R., & Becker, B. (1986). A scale for the assessment of object relations: reliability, validity, and factorial invariance. *Journal of Clinical Psychology*,*42*, 733-741.

Bem, D., & Honorton, C. (1994). Does Psi exist? Replicable evidence for an anomalous process of information transfer. *Psychological Bulletin, 115*, 4-27.

Bevis, J. (1986). Connectedness versus separateness: understanding male/female differences in self and relationship. (Unpublished doctoral dissertation, Boston University).

Blatt, S., & Ritzler, B. (1974). Thought disorder and boundary disturbance in psychosis. *Journal of Consulting and Clinical Psychology, 42*, 370-381.

Broadbent, D. (1971). William James Lectures. In Broadbent, D. (1973) *In Defense of Empirical Psychology*. London: Methuen.

Costa, P., & McCrae, R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) Professional Manual*. Odessa, FL: Psychological Assessment Resources.

Cowen, D. & Levin, R. (1995). The use of the Hartmann Boundary Questionnaire with an adolescent population. *Dreaming, 5*, 105-114.

Earle, J. (1992). Social desirability and thin boundaries. Unpublished manuscript.

Ehrman, M., & Oxford, R. (1995). Cognition plus: correlates of language learning success. *The Modern Language Journal, 79*, 67-89.

Federn, P. (1952). *Ego psychology & the psychoses*. New York: Basic Books.

Fisher, S. (1992). Personal Communications.

Fisher, S., & Cleveland, S. (1968). *Body image & personality* (2nd ed.). New York: Dover.

Freud, S. (1923). The ego and the id. In *The standard edition of the complete psychological works of Sigmund Freud*, vol. 19 (pp. 3-66). London: Hogarth Press.

Funkhauser, A, Würmle, O., Cornu, C., & Bahro, M. (2001). Dream life & intrapsychic boundaries in the elderly. *Dreaming, 11*, 83-88.

Galvin, F. (1993). The effect of lucid dream training upon the frequency & severity of nightmares Unpublished Doctoral dissertation, Boston University.

Gardner, R., Holzman, P., Klein, G., Linton, H., & Spence, D. (1959). Cognitive control: a study of individual consistencies in cognitive behavior. *Psychological Issues*, Monograph 4. New York: International Universities Press.

Garg, M., & Hartmann, E. (1993). Neuropsychology of persons with thick or thin boundaries. Unpublished manuscript.

Gsell, P. (1971). *Rodin on Art*. New York: Horizon Press.

Harrison, R., Hartmann, E., & Bevis, J. (1989). The Hartmann Boundary Questionnaire: a measure of thin & thick boundaries. Manuscript submitted for publication.

Hartmann, E. (1989). Boundaries of dreams, boundaries of dreamers: thin & thick boundaries as a new personality dimension. *Psychiaric Journal of the University of Ottawa, 14*, 557-560.

Hartmann, E. (1991). *Boundaries in the Mind*. New York: Basic Books.

Hartmann, E. (1992). Boundaries in the mind: boundary structure related to sleep & sleep disorders. *Sleep Research, 21*, 126.

Hartmann, E. (1996). Outline for a theory on the nature & functions of dreaming. *Dreaming, 6*, 265-271.

Hartmann, E. (1997). The concept of boundaries in counseling & psychotherapy. *British Journal of Guidance & Counseling, 25*, 147-162.

Hartmann, E. (1998). *Dreams & Nightmares*. New York: Plenum Press (1998), Perseus (2000).

Hartmann, E., Baekel, F., & Zwilling, G. (1972). Psychological differences between long & short sleepers. *Arch Gen Psychiat*, *26*, 463-468.

Hartmann, E., Harrison, R., Bevis, J., Hurwitz, I., Holevas, A., & Dawani, H. (1987). The Boundary Questionnaire: a measure of thin & thick boundaries derived from work with nightmare sufferers. *Sleep Research*, *16*, 274.

Hartmann, E. Elkin, R., & Garg, M.(1991). Personality & Dreaming: the dreams of people with very thick or very thin boundaries. *Dreaming*, *1*, 311-324.

Hartmann, E., Rosen, R., Grace, N. (1998). Contextualizing images in dreams & daydreams. *Sleep*, *21S*, 279.

Hartmann, E., Rosen, R. (1998). Personality & dreaming: boundary structure & dreams. *Dreaming*, *8*, 31-40.

Hartmann, E., & Zborowski, M. (2001). Dreams: correlates of the Contextualizing Image. *Sleep*, *24S*, A174.

Hartmann, E., Zborowski, M., McNamara, P., Rosen, R., & Grace, N. (1999). Contextualizing images in dreams: relationship to the emotional state of the dreamer. *Sleep*, *22S*, 131.

Hartmann, E., Zborowski, M., Rosen, R., & Grace, N. (2001). Contextualizing images in dreams: more intense after abuse & trauma. *Dreaming*, 11, 115-126.

James, W. (1907). *Pragmatism: A New Name for Some Old Ways of Thinking*. New York: Washington Square Press, 1983.

Jawer, M. (2001). *The Emotional Gateway: The Dynamics of Feeling & their Link to Paranormal Perception*. Manuscript submitted for publication.

Krippner, S., Wickramasekera, I., Wickramasekera, J., & Winstead, C. (1998). The Ramtha phenomenon: psychological, phenomenological, & geomagnetic data. *The Journal of the American Society for Psychical Research*, *92*, 1-24.

Kunzendorf, R., & Maurer, J. (1988-89). Hypnotic attenutation of the 'boundaries' between emotional, visual, & auditory sensations. *Imagination, Cognition & Personality*, *8*(3), 225-234.

Kunzendorf, R., Hartmann, E., Cohen, R., & Cutler, J. (1997). Bizarreness of the dreams and daydreams reported by individuals with thin and thick boundaries. *Dreaming*, *7*, 265-271.

Landis, B. (1970). Ego boundaries. *Psychological Issues*, *6*(4), Monograph #24. New York: International Universities Press.

Lange, R., Thalbourne, M., Houran, J., & Storm, L. (2000). The revised Transliminality Scale: reliability and validity data from a Rasch top-down purification procedure. *Consciousness and Cognition*, *9*, 591-617.

Levin, R., Galen, J. & Zywiak, B. (1991). Nightmares, Boundaries, and Creativity. *Dreaming* 1, 63-74.

Levin, R., & Fireman, G. (1993). Psychoanalytic studies on the experience of schizophrenia. In Masling, J. & Bornstein, R., (Eds.), *Psychoanalytic Perspectives on Psychopathology* (pp. 139-189). Washington, D.C.: APA Press.

Levin,R., Gilmartin, L., & Lamontanaro, L. (1998-99). Cognitive style and perception: the relationship of boundary thinness to visual-spatial processing in dreaming and waking thought. *Imagination, cognition, and personality, 18*(1), 25-41.

Lewin, K. (1936). *Principles of topological psychology*. New York: McGraw-Hill.

McCrae, R. (1994). Openness to experience: expanding the boundaries of factor V. *European Journal of Personality, 8*, 251-272.

Mednick, S. (1962). The associative basis of the creative process. *Psychological Review, 69*, 220-232.

Rader, C., Kunzendorf, R., & Carrabino, C. (1996). The relation of imagery vividness, absorption, reality boundaries and synesthesia to hypnotic states and traits. In: *Hypnosis and Imagination*, Kunzendorf R., Spanos, N., Wallace, B., (eds). New York: Baywood Publishing Company. pp. 99-121.

Rokeach, M. (1960). *The Open and Closed Mind.* New York: Basic Books.

Ryan, C. (2000). Personality of fashion models. (Unpublished master's thesis, Saybrook Graduate School, San Francisco).

Sand, S. & Levin, R. (1996). Concordance between Hartmann's Boundary Questionnaire and the Eysenck Personality Inventory. *Perceptual and Motor Skills, 82*, 192-194.

Schredl, M., Kleinferchner, P., & Gell, T. (1996). Dreaming and personality: thick vs. thin boundaries. *Dreaming, 6*, 219-223.

Schredl, M., Schäfer, G., Hofmann, F., & Jacob, S. (1999). Dream content and personality: thick vs. thin boundaries. *Dreaming, 9*, 257.

Spence, D. (1964). Conscious and preconscious influences on recall: another example of the restricting effects of awareness. *Journal of Abnormal and Social Psychology, 68*, 92-99.

Thalbourne, M. (1991). The psychology of mystical experience. *Exceptional Human Experience, 9*, 168-186.

Thalbourne, M. (1996). Belief in life after death: psychological origins and influences. *Personality and Individual Differences, 21*, 1043-1045.

Watson, R. (July, 1985). Phasic integrated potentials and ego boundary deficit Paper presented to a joint meeting of the Sleep Research Society and the Association of Sleep Disorders Centers. Seattle, Washington.

Zborowski, M., McNamara, P., Hartmann, E., Murphy, M., & Mattle, L. (1998). Boundary structure related to sleep measures and to dream content. *Sleep, 21S*, 284.

Zborowski, M., Hartmann, E., & Newsom, M. (2001). The Hartmann Boundary Questionnaire: two studies examining personality correlates and interpersonal behavior. Submitted for publication.

Note: This paper was sponsored by James Houran, Southern Illinois University School of Medicine, Springfield, IL.

Effect of Humor on Hitchhiking: A Field Experiment

Nicolas Guéguen
Université de Bretagne-Sud - Vannes

Studies have shown that a positive mood, activated by an amusing drawing, leads to increased helping behavior. Four hitchhiking confederates, 2 young men and 2 young women, tested 1600 motorists. Each confederate held a sign indicating the place where he/she wished to go. In half of the cases, a funny drawing (a smiling face) appeared on the sign. Contrary to my expectations, the funny drawing inhibited the helping behavior of the motorists, but only when the hitchhiker was a woman.

Multiple studies have shown that a positive mood is more conducive to helping behavior than a neutral mood (Bizman, Yinin, Ronco & Schachar, 1980; Forgas, 1997; 1998; Harris & Smith, 1975; Job, 1987; Levin & Isen, 1975; Rind, 1997; Weyant, 1978). This positive mood can be activated in many different ways and usually does not require elaborate tools. A false attribution of success or failure in a task activates a positive/negative mood, which, in turn, affects altruism in a way which is congruent with the mood (Clark & Waddell, 1986). Finding a coin in a phone booth or on the ground is enough to induce a positive mood, which in return enhances altruism (Batson, Coke, Chard, Smith & Talaferro, 1979; Blevins & Murphy, 1974; Kidd & Marshall, 1982 Isen & Levin, 1972; Isen & Simmonds, 1978; Levin & Isen, 1975). To be offered candy or a cookie produces the same effect (Harris & Smith, 1975; Isen & Levin, 1972).

Other environmental factors also activate a positive mood, which affects pro-social behavior. For example, the amount of money in tips given by customers increases on sunny days (Cunningham, 1979). A study by Rind (1996) showed that information about the weather, given by the waiter to customers who had not yet seen the color of the sky, was sufficient to make his tip vary. Customers left the waiter a large or a small tip depending on whether he told them it was sunny or rainy. Pleasant ambient smells (e.g. baking, cooking, roasting coffee) in a shopping mall led the passers-by to provide change for a dollar to a same-sex confederate more readily than in the absence of such odors (Baron, 1997). Pleasant music also plays an important part in helping behavior.

Author info: Correspondence should be sent to: Dr. Nicolas Guéguen, Laboratoire GRESICO, Université de Bretagne-Sud, IUT de Vannes-Département TC. 8, rue Montaigne. BP561 - 56017 VANNES - France.

North American Journal of Psychology, 2001, Vol. 3, No. 2, 369-376.

© NAJP

BEST COPY AVAILABLE

Fried and Berkowitz (1979) found that subjects who heard soothing music for seven minutes were more apt to accept a request for help made immediately afterwards by the experimenter than subjects in a no-music condition. Another way of inducing a mood state in a subject consists of exposing him or her to pleasant versus unpleasant pictures or texts (Forgas, 1997; 1998). For example, drawing a smiling face on the bill increased the amount of tip money given to a waitress (Rind & Bordia, 1996). In the same way, a hand-made drawing of the sun ("smiley-face"), added to the bottom of the bill of customers having a drink, led them to leave a larger tip for the waiters or waitresses (Guéguen & LeGohérel, 2000). Attaching a small card to the bill onto which a joke was appended produced the same effect (Guéguen, in press).

In the present study a new evaluation of the effect of a drawing was made in conjunction with one of the standard requests for help: hitchhiking. It is known that, in this context, traditional factors of influence lead to increased helping behavior. For instance, Snyder, Grether, and Keller (1974) and Morgan, Lockard, Fahrenbrugh and Smith (1975) have shown that motorists are more likely to stop if hitchhikers look at them straight in the face than if they look somewhere else. The effect of eye contact on various helping behaviors has been observed in several studies (Brockner, Pressman, Cabitt, & Moran, 1982; Bull, & Gibson-Robinson, 1981; Hornik & Ellis, 1988; Kleinke, 1980). Crassweller, Gordon and Tedford (1972) also found that a male hitchhiker is more likely to get help when he is well-dressed. Similarly, several studies have shown that people were more likely to respond to a request for help made by solicitors who were well-dressed (Goodman & Gareis, 1993; Kleinke, 1977; McElroy & Morrow, 1994). Pomazal and Clore (1973) found that drivers were more likely to help a hitchhiker in a high need situation (e.g.: when the hitchhiker confederate had a physical disability). So, in keeping with Rind and Bordia's (1996) study, I expected a "smiley" face drawing placed on the sign indicating the hitchhiker's destination, would increase the rate of drivers who would stop to help the solicitor.

Because research on hitchhiking has shown that females receive more rides than males (Clifford & Cleary, 1971; Pomazal & Clore, 1973; Morgan et al., 1975; Snyder et al., 1974), it was expected than female hitchhikers would receive more help in the present experiment.

METHOD

Participants

Participants were 1600 drivers, men and women, solicited at the entry of a famous peninsula ("Presqu'Île de Rhuys") of Brittany in France. The peninsula is about 24 kilometers long.

Procedure

Four people, 2 men and 2 women, 19-21 years old, served as confederates in this experiment. All of them were first year students from the Department of Business at the University of Bretagne-Sud in France, and all volunteered to participate as confederates in this experiment. Both female and male confederates were selected by 2 other male and 2 other female evaluators who were asked to select confederates based on their physical attractiveness. These evaluators were instructed to carefully examine photographs of students and to select people, males and females, who were average in attractiveness. After that, these students were asked to participated in the experiment. All of them accepted the solicitation of the experimenter. One confederate stood at the side of the road at a place with good visibility for the motorists and with a broad berm, making the stopping and restarting of vehicles quite easy and safe. The experiment took place between 2 p.m. and 4 p.m. during weekends on particularly sunny and pleasant days at the beginning of summer. The confederates were dressed in a casual way for young people of their age (t-shirts, jeans, and tennis shoes of light colors). Each confederate was standing alone on the side of the road, holding in his/her hand a paperboard sign of 30 cm × 20 cm on which was written the name of a city located halfway up of the peninsula (approximately 11 kilometers). According to the condition, a smiling face (☺) was (or was not) drawn on the upper left side of the sign. A pretest had shown that the drivers could see the drawing on the sign. The confederates were instructed to always hold the sign at the same height and not to look at the drivers in the eyes, but to look at an unknown point opposite them. If a driver stopped, the confederate was then instructed to debrief the participant. He/she explained to the driver that he/she was conducted an experiment on hitchhiking. Then, the driver was warmly thanked for his help. Once 40 cars had passed by (7.13 minutes waiting on average, *SD* = 4.46) the confederate changed signs. When each confederate had tested 400 drivers, the experiment stopped.

RESULTS

All the drivers who stopped were men, rendering impossible any comparison of the sex of the drivers. The number of drivers who stopped is presented in Table 1, categorized according to the confederate's gender and the experimental condition. Since no differences between the confederates of the same gender were found between the two men and the two women, the data were aggregated.

A main effect of the confederates' gender was obtained. Generally speaking, female confederates were more likely to get a ride than male hitchhikers (X^2 (1, N = 1600) = 14.43, $p < .001$). A main effect of the type of sign used was also observed (X^2 (1, N = 1600) = 7.59, $p < .01$). However, contrary to my assumption, the presence of the smiley face was associated with a decrease in the number of drivers who offered to give a lift. A significant interaction, tested by way of a 2 (male and female confederate) × 2 (no drawing/smiley face) × 2 (driver stop/no stop) log-linear analysis, between the confederates' gender and the type of sign used was observed (X^2 (4, N = 1600) = 23.59 $p < .001$). The presence of the smiley face did not have any effect at all on drivers' stopping rates when the hitchhiker was a man (X^2 (1, N = 800) = 0.33, *ns*). In contrast, the smiley face led to decreased stopping rates when the hitchhiker was a woman (X^2 (1, N = 800) = 6.82, $p < .01$).

TABLE 1: Number of Participants who Stopped According to the Type of Sign Used and the Gender of the Confederate[1]

	Sign Without Drawing	Sign With Smiling Face
Male Confederates	29 (7.3%)	24 (6.0%)
Female Confederates	62 (15.5%)	35 (8.8%)

[1] There were 400 drivers per case

DISCUSSION

Results of the present study showed that, in general, female hitchhikers were helped more than male hitchhikers. This finding is consistent with other field experiments conducted on hitchhiking (Clifford & Cleary, 1971; Morgan et al., 1975; Snyder et al., 1974). In other studies more help was offered to female hitchhikers even in high need situations. Pomazal and Clore (1973) have shown that when the hitchhiker-confederate presented a physical disability, females were helped more than males. Such results are also consistent with the literature of helping behavior in general. When the helper is a male, a woman in apparent need tends to be helped more than a man in apparent need (Bickman, 1974; Dovidio, 1982; Gore, Tobiasen & Kayson, 1997; Latané & Dabbs, 1975; Simon, 1976). This effect is also accentuated in a high need situation (Juni & Rooth, 1981).

Contrary to our hypothesis, the presence of the smiling face on the sign decreased the number of drivers stopping and offering rides to the hitchhiking confederates. However this phenomenon was observed only when the hitchhiker was a woman, whereas no difference between the two types of sign was observed when the hitchhiker was a man. These results contrast sharply with those of previous studies. Rind and Bordia

(1996) have shown that when the "smiley" face appeared on the back of customers' bill, it led patrons to give more tip money to a waitress. When a waiter drew a "smiley" face no difference was found. In the present experiment, when the drawing appeared on the sign of the two male confederates it had no effect on the drivers stopping rate, just as it had no effect on tipping in Rind and Bordia's study. Yet, this drawing apparently affected the drivers negatively in the present study when the hitchhiker was a woman. Of course, my research does not correspond to that carried out by Rind and Bordia (1996) on several points. The dependent variable was different, the subjects of Rind and Bordia were in groups of both men and women, whereas my participants were primarily men driving alone, as was true in other field studies conducted on hitchhiking (Forsyth, 1978; Pomazal & Clore, 1973). The two studies were conducted in different countries (USA/France). However, the cultural difference seems an unlikely explanation, since Guéguen and LeGohérel (2000), and Guéguen (In press), showed that, in France, bar customers reacted positively (i.e. they left larger tips to waiters and waitresses) to a funny drawing appearing on the bill or to a joke written on a card accompanying the bill. In both studies, no difference was observed according to the server's gender.

The effect observed may come from a differential reaction to the drawing when it comes from a woman. Studies have shown that people react differently to the same joke according to whether it is told by a man or a woman (Herzog, 1999; Mundorf, Bhatia, Zillman, Lester & Robertson, 1988; Van Giffen & Maher, 1995). It is possible that, in the present study the drawing led the drivers to perceive the female confederate less positively, which, in return, resulted in decreasing their likelihood of helping. It may also be that this decrease is due to a perception of incongruity between the smiling face and the hitchhiker's gender.

Perhaps humor on the part of a female-hitchhiker was perceived by male drivers as a sign of less femininity because being humorous is traditionally perceived as a male characteristic (Wilson & Molleston, 1981). Harris and Bays (1973) have shown that women solicitors were helped more when wearing feminine than when wearing masculine attire. Tice and Baumeister (1985) have also found such results in an emergency situation. In my experiment, humor may had had the same effect.

Of course, the results of the present study provide no proof for such explanations. In order to evaluate them, further experiments need to be conducted. The "less positive perception" hypothesis created by a woman hitchhiker holding a smiling face on her sign should be evaluated. In the two experimental conditions, photographs of the confederates could be used as a stimulus to evaluate the impression created by the hitchhiker and the probability of stopping in such conditions. The appropriateness of humor in accordance with the confederate's gender needs to be evaluated in the same way.

REFERENCES

Baron, R. (1997). The sweet smell of... helping: Effects of pleasant ambient fragance on prosocial behavior in shopping mall. *Personality and Social Psychology Bulletin, 23*, 498-503.

Batson, D., Coke, J., Chard, F., Smith, D., & Talaferro, A. (1979). Generality of the " glow of goodwill ": Effects of mood on helping and information acquisition. *Social Psychology Quarterly, 42*, 176-179.

Bickman, L. (1974). Sex and helping behavior. *The Journal of Social Psychology, 92*, 43-53.

Bizman, A., Yinin, Y., Ronco, B., & Schachar, T. (1980). Regaining self-esteem through helping behavior. *The Journal of Psychology, 105*, 203-209.

Blevins, G., & Murphy, T. (1974). Feeling good and helping: Further phone booth findings, *Psychological Reports, 34*, 326.

Brockner, J., Pressman, B., Cabitt, J., & Moran, P. (1982). Nonverbal intimacy, sex, and compliance: A field study. *Journal of Nonverbal Behavior,6*, 253-258.

Bull, R., & Gibson-Robinson, E. (1981). The influence of eye-gaze, style of dress, and locality on the amounts of money donated to a charity. *Human Relations, 34*, 895-905.

Clark, M., & Waddell, B. (1986). Effects of mood on thoughts about helping, attraction and information acquisition. *Social Psychology Quarterly, 56*, 31-35.

Clifford, M., & Cleary, P. (1971). The odds in hitchhiking. *Psychology Today, 5*, 38.

Crassweller, P., Gordon, M., & Tedford,W. (1972). An experimental investigation of hitchhiking. *The Journal of Psychology, 82*, 43-47.

Cunningham, M. R. (1979). Weather, mood and helping behavior: Quasi experiments with the sunshine Samaritan. *Journal of Personality and Social Psychology, 37*, 1947-1956.

Dovidio, J. (1982). Sex, costs, and helping behavior. *The Journal of Social Psychology, 112*, 231-236.

Forgas, J. (1997). Affect and strategic communication: The effects of mood on the production and interpretation of verbal requests. *Polish Psychological Bulletin, 28*, 145-173.

Forgas, J. (1998). Asking nicely? The effects of mood on responding to more or less polite request. *Personality and Social Psychology Bulletin, 24*, 173-185.

Forsyth, S. (1978). Predicting motorists' altruism. *Psychological Reports, 43*, 567-572.

Fried, R., & Berkowitz, L. (1979). Music hath charms... and can influence helpfulness. *Journal of Applied Social Psychology, 9*, 199-208.

Goodman, M., & Gareis, K. (1993). The influence of status in the decision to help. *The Journal of Social Psychology, 133*, 23-31.

Gore, K., Tobiasen, M. & Kayson, W. (1997). Effects of sex of caller, implied sexual orientation of caller, and urgency on altruistic response using the wrong number technique. *Psychological Reports, 80*, 927-930.

Guéguen, N., & LeGohérel, P. (2000). Effect on barmen's tipping of drawing a sun on the bottom of customers' checks. *Psychological Reports, 87*, 223-226.

Guéguen, N. (In Press). The effects of a joke on tipping when it is delivered at the same time as the bill. *Journal of Applied Social Psychology.*

Harris, M., & Bays, G. (1973). Altruism and sex roles. *Psychological Reports, 32*, 1002.

Harris, M., & Smith, R. (1975). Mood and helping. *The Journal of Social Psychology, 91*, 215-221.

Herzog, T. (1999). Gender differences in humor appreciation revisited. *Humor: International Journal of Humor Research, 12*, 411-423.

Hornik, J., & Ellis, S. (1988). Strategies to secure compliance for a mall intercept interview. *Public Opinion Quarterly, 52*, 539-551.

Isen, A., & Levin, P. (1972). Effect of feeling good on helping. *Journal of Personality and Social Psychology, 21*, 384-388.

Isen, A., & Simmonds, S. (1978). The effect of feeling good on a helping task that is incompatible with good mood. *Social Psychology Quarterly, 41*, 346-349.

Job, S. (1987). The effect of mood on helping behavior. *The Journal of Social Psychology, 127*, 323-328.

Juni, S., & Roth, M., (1981). Sexism and handicapism in interpersonal helping. *The Journal of Social Psychology, 115*, 175-181.

Kleinke, C. (1977). Effects of dress on compliance to requests in a field setting. *The Journal of Social Psychology, 101*, 223-240.

Kleinke, C. (1980). Interaction between gaze and legitimacy of request on compliance in a field setting. *Journal of Nonverbal Behavior, 5*, 3-12.

Kidd, R., & Marshall, L. (1982). Self-reflection, mood, and helpful behavior. *Journal of Research in Personality, 16*, 319-334.

Latané, B. & Dabbs, J. (1975). Sex, group size and helping in three cities. *Sociometry, 38*, 180-194.

Levin, P., & Isen, A. (1975). Further studies on the effect of feeling good on helping. *Sociometry, 38*, 141-147.

McElroy, J., & Morrow, P. (1994). Personal space, personal appearance, and personal selling. *Psychological Reports, 74*, 425-426.

Morgan, C., Lockard, J., Farenbrugh, C., & Smith, J. (1975). Hitchhiking: Social signals at a distance. *Bulletin of the Psychonomic Society, 5*, 459-461.

Mundorf, N., Bhatia, A., Zillman, D., Lester, P., & Robertson, S. (1988). Gender differences in humor appreciation. *Humor: International Journal of Humor Research, 1*, 231-243.

Pomazal, R., & Clore, G. (1973). Helping on the highway: The effects of dependency and sex. *Journal of Applied Social Psychology, 3*, 150-164.

Rind, B. (1996). Effect of beliefs about weather conditions on tipping. *Journal of Applied Social Psychology, 26*, 137-147.

Rind, B. (1997). Effect of interest arousal on compliance with a request for help. *Basic and Applied Social Psychology, 19*, 49-59.

Rind, B., & Bordia, P. (1996). Effect on restaurant tipping of male and female servers drawing a happy, smiling face on backs of customers' checks. *Journal of Applied Social Psychology, 26*, 218-225.

Simon, A. (1976). Chivalry on the road: Helping stalled drivers. *Psychological Reports, 39*, 883-886.

Snyder, M., Grether, J., & Keller, K. (1974). Staring and compliance: A field experiment on hitchhiking. *Journal of Applied Social Psychology, 4*, 165-170.

Tice, D., & Baumeister, R. (1985). Masculinity inhibits helping in emergencies: Personality does predict the bystander effect. *Journal of Personality and Social Psychology.* 2, 420-428.

Van Giffen, K., & Maher, K. (1995). Memorable humorous incidents: Gender, themes, and setting effects. *Humor: International Journal of Humor Research, 8*, 39-50.

Weyant, J. (1978). Effects of mood states, costs, and benefits of helping. *Journal of Personality and Social Psychology, 36*, 1169-1176.

Wilson, D., & Molleston, J. (1981). Effects of sex and type of humor on humor appreciation. *Journal of Personality Assessment, 45*, 90-96.

Religion's Role in the Terrorist Attack of September 11, 2001

Michael E. Nielsen
Georgia Southern University

The central role of religion in the terrorist attacks on New York and Washington, D.C., is addressed. The terrorists were motivated by a militant fundamentalist worldview, which rejects modernism. According to Osama bin Laden, corrupt Western values and mores are not to be tolerated, making attacks justified. Religion's role in coping with the attacks also is described using Pargament's theoretical framework. Steps that individuals can take to deal constructively with the situation are listed.

Four planes were hijacked on Tuesday, September 11, 2001. Two of the planes were crashed deliberately into New York City's World Trade Center towers, causing their collapse. A third plane was flown into the Pentagon, outside Washington DC. The fourth plane, after a struggle between hijackers and passengers, crashed in a field before hitting its target. An untold number of people are dead as a result of these actions, which also are expected to affect people throughout the world as its repercussions are felt in the economic and political spheres. Clearly, this coordinated, concerted attack was intended to bring down symbols of US economic and military might. Early reports also indicate that important centers of US political power - the White House and the president's airplane - also were targets.

Militant Fundamentalist Worldview

Evidence currently points to militant Islamic fundamentalists affiliated with Osama bin Laden. One may expect that their actions were motivated by several factors, but we are left with an incomplete understanding of this group without an effort to grasp the central role of religion in their lives and how it influences their perception of the world.

Osama bin Laden, a tall man in his mid-forties, grew up in Saudi Arabia, where his father became rich in the building industry. From his father's inheritance bin Laden gained $250 million, which he has used to finance and support his view of Islam. The Sunni branch of Islam is most common in Arabia and in the world, and bin Laden adheres to a radical

---

Author info: Correspondence should be sent to: Dr. Michael Nielsen, Department of Psychology, Georgia Southern University, Statesboro, GA 30460-8041, or at mnielsen@gasou.edu.

North American Journal of Psychology, 2001, Vol. 3, No. 3, 377-384.

© NAJP

fundamentalist interpretation of it. Like all of Islam, the Sunnis believe in one God, and that Muhammed was God's last prophet. Islam encourages people to be devoted to God, but also to be generous, hospitable, and to demonstrate care for others. So, the important questions for us to understand regarding the terrorist attack are "What makes something fundamentalist?" and "How did this worldview influence the terrorist actions?"

The overarching theme to the fundamentalist, whether Muslim or Christian, is that God is to be worshipped, respected, feared & obeyed above all else. All other considerations take a back seat to God. This intense and abiding devotion means that there are some things that are completely, utterly non-negotiable. In this way, it is like viewing the world as being black and white, with little if any gray between that which is good and that which is evil. Accompanying this is a tendency toward literalism. If scripture says that Noah built an ark, put two of each type of animal throughout the world on it, and that they sailed on that boat while the entire earth was flooded, then it happened. No questions need be asked; it happened, regardless of whether it is logically consistent with what we know about animals, floods, ancient ships or the geological record.

An important element of the fundamentalist mindset is the rejection of modernism. Contemporary western values are inconsistent with God's values and with His will for humanity. People's duty is to worship God, not to ignore or ridicule Him. God's rules are clear and they are to be enforced and respected, not flaunted. As a result, fundamentalists have very conservative views on social issues. The fundamentalist finds media portrayals of sex, inappropriate gender roles, and many other elements of Western culture to be abominations completely out of harmony with God's will. God does not like it, doesn't tolerate it, and neither do God's devoted followers.

Bin Laden

Bin Laden was exiled from Saudi Arabia for his views in 1991, during the Gulf War. Bin Laden objected to the Saudi government policy that allowed infidels, US and other non-Muslim military troops, on their sacred soil (Saudi Arabia is considered sacred because the most holy of Islam's sites are located there.). By exiling bin Laden, the Saudi government avoided complications within its borders and with its allies. Bin Laden went to Sudan, where he used his inheritance to help the country's poor (by developing jobs, etc.), and also is said to have continued his efforts against the west. In 1996, Sudan bowed to international pressure and exiled him.

From Sudan, bin Laden went to Afghanistan where he has stayed until the present time. For bin Laden, this was a return to Afghanistan. He went there earlier, in the mid-1980s, during the country's ten-year war with the Soviet Union. In 1989 the Soviet Union left, defeated, and an intense civil war raked the country. During that ongoing war, a group known as the Taliban gained control of the majority of Afghanistan; a separatist faction remains in northern Afghanistan. The Taliban, who share a deeply rigid, radical fundamentalist view of Islam, are sympathetic to bin Laden, and have allowed him to maintain training camps to support members of his group al Qaeda. It also is important to note that most countries never recognized the Taliban's government. Indeed, the civil war continues.

Bin Laden formed the group al Qaeda (Arabic for "The Base") as a vehicle to help him achieve his goals for Islam. The consistent theme of the group has been to fight against the West, most chiefly against the US, but also against countries that aid US interests. Among other things the group has been implicated in several terrorist attacks, perhaps most notably the 1998 US Embassy bombings in Kenya and Tanzania which killed 224 people. Frequently these attacks have come after a declaration (fatwa) denouncing the US and its socio-political agenda. For example, a May 1998 declaration entitled "The Nuclear Bomb of Islam" included the statement that "it is the duty of Muslims to prepare as much force as possible to terrorize the enemies of God" (Washington File, 1999). A few months later, the embassies were bombed; the US responded with cruise missile attacks in Afghanistan (bin Laden's suspected training camp) and Sudan (where he stored or made the weapons believed to have been used in the attack).

It is worth mentioning that in August, 1996, bin Laden issued a declaration of jihad. Its lengthy title makes clear bin Laden's position: "Message from Osama bin Laden to his Muslim Brothers in the Whole World and Especially in the Arabian Peninsula: Declaration of Jihad against the Americans Occupying the Land of the Two Holy Mosques; Expel the Heretics from the Arabian Peninsula" (Washington File, 1999). This declaration of jihad reflects the meaning that most westerners have come to associate with the word. Jihad, however, has been called the most misunderstood word in religion today. Although we may read of it in this context, its common, usual meaning is that of a struggle to purify oneself. In many ways, it is more appropriate to consider jihad as having the connotation similar to that of a fast, or going on a retreat, to a Christian.

People who view bin Laden as a typical Muslim, make a serious mistake, as do those who view any single individual as a prototypic example of any group. For example, a selection of public figures who profess Christianity might include: Jerry Falwell, Pope John Paul II, David Koresh, Bill Clinton, and George W. Bush. Each claim Christianity, but few would consider each of these men to represent Christianity equally well. More broadly, Aryan Nation groups that use Christianity to further their social cause of racial separation and hatred are not considered by most people to be prototypic Christians. Likewise, bin Laden is not a typical Muslim, and militant fundamentalist Muslims are not representative of Islam.

Regardless of whether it was bin Laden or someone else who coordinated the attacks, there are more general questions for us to consider. Why are we surprised when people act in accordance with their religious beliefs? This certainly isn't the first time it has happened; examples abound. In 1997, a group of 39 people known as Heavens Gate committed suicide. "Heaven is a beautiful place," they said, and they wanted to go there. It shocked the country.

During World War II, some Japanese pilots went on Kamikaze missions. Many people learned with disbelief that these pilots, with just enough fuel to reach their targets, flew directly into the enemy and sacrificed themselves in order to serve the emperor, their God.

This was not the first time that extreme views have resulted in extreme behavior. Nor will it be the last, because attitudes do affect behavior.

In light of this, it is instructive to ask about attitudes and behavior regarding the attack. How do individuals use religion to cope with disasters such as this?

Religious Coping

Ken Pargament (1997) describes many different ways that religion helps people deal with traumatic, life-changing events. Examining his theory will help us make some sense of people's reactions to the attack. The basic assumption underlying Pargament's theory is that people search for significance in their lives. We are goal-directed, and try to find and hold on to the things that give our life a sense of significance and meaning. For many people, religion provides a way to do this. It gives us something to strive for, and tells us how we can obtain it.

Crises, however, threaten one's sense of significance (Solomon, Greenberg & Pyszczynski, 2000). They put at risk the things that are valued, and many people use religion in order to regain and maintain that which is significant. It does this in two basic ways. The first method, which Pargament calls *conservation*, maintains the same items as being significant. People prefer conservational strategies because they involve less change, less stress. The alternative, *transformation*, is used when people determine that their current sense of significance is untenable, that they must change their source of significance.

Conservational coping comes in four general types.

Religious prevention helps us maintain the sense of significance by declaring certain things as being off limits. Drug use and extra-marital sex, for example, may be determined to be out of the bounds of acceptable behavior. Avoiding these activities helps people maintain an ongoing sense of purpose, that they are headed in the right direction, so to speak. Although it often refers to reduced rates of behavior that put one at risk, comments from Jerry Falwell and Pat Robertson (Harris, 2001), who suggested that overly permissive societal norms are the root cause of the terrorist attack, also illustrate this perspective.

Religious support maintains the sense of significance by helping people feel as though they are not alone in their struggle or stress. God may be viewed as a helping partner, and a group of like-minded people may lend a sense that they are receiving help during a time of need. Many people's comments, such as "God is with us," reflect this form of religious coping.

Ritual purification offers a way for people to address personal failures and inadequacies. By participating in purification rites such as confession or repentance individuals can reconcile their ideal, their sense of significance, with the reality that they don't always measure up to those ideals. The countless prayer services and candle-light vigils that have been performed, and the millions of people who have participated in them, demonstrate the importance of ritual purification as a coping method to people distressed by the terrorist acts.

Religious reframing enables the individual to reinterpret (or "reframe") things so that the sense of meaning and significance is maintained. This occurs, for example, if the negative events take on a new meaning, and perhaps even become an opportunity to grow. Many of the messages heard from political leaders, such as the idea that terrorists and bombs can not destroy freedom, fit into this category. They reaffirm core values and identities, and in so doing maintain individuals' views of themselves and what they consider to be important.

Transformational coping strategies are more difficult to deal with because they require more change. If for some reason individuals cannot maintain their sense of significance, if the things that they value actually lose value, then people will opt for transformational coping. An example of this is a conversion or apostasy experience. Suppose that, after learning of the attacks and the tremendous loss of life, a person lost faith in God and humanity, forsaking them for a new belief system. This would be a transformational coping strategy. Changing one's belief system is a major undertaking, and typically comes only reluctantly, after conservational coping strategies have been tried and failed. Recent media reports have not yet reflected this type of coping but because people first attempt conservational strategies, we may expect such accounts in the future.

Religion provides a powerful way to cope with stress. Pargament's (1997) research shows that religion adds additional resources to the other coping resources that a person has available. Even if one has a close-knit family, the feeling that one is part of a religious group helps the individual to weather difficult times. Also, because religion is abstract and symbolic, one can reinterpret events so that they make sense within the core belief system. It lends itself to reframing.

The Future

Where do we go from here? Clearly, typical citizens are not empowered with the ability to strike back at the terrorists, whoever they are. Nevertheless, there are some things that people may do to benefit themselves and their communities.

Individuals can reach out to those in need. Many reasonable persons feel that we have a moral obligation to offer help to those who suffered personally in the attacks. There also are many legitimate relief agencies that can use your time or monetary contribution. Several are listed at http://www.psyww.com/psyrelig/refugee.htm.

Talk about the attacks, about their implications for daily life. There will be debates about the nature of freedom and democracy, and what freedoms people might be willing to forego for the hope of greater safety. Think about these important issues and discuss them with other people. In all likelihood these debates are going to become increasingly important in the future.

Individuals can be involved in the political process where they live by writing to their elected officials, urging them to use care in responding to these heinous acts. For example, some commentators in the US and in the world remind us that much of the world sees the US as being hypocritical for preaching the virtue of democracy, at the same time that it refuses to encourage elections in countries led by what amounts to dictatorships. Do people in the US believe deeply in democracy, or do short-term expediencies win out? These are questions that American citizens should be debating. Discuss these issues, and write to your elected officials, urging them to do the same. Ask what the US can do to help its actions match its lofty ideals.

Finally, people can learn about Islam and share that knowledge with other persons. One very good source of information about Islam is http://www.religioustolerance.org. Treat Muslims well, and refuse to tolerate or condone acts of abuse. During World War II, the US held captive many of its citizens simply because of their race. The US, as a society, has learned something during the 60 years since that war. Let us show each other and the world that we reject prejudice and hate.

William James (1902/1985) suggests that we can judge religion by its effects. If religion helps people deal with tragedy, then that is good, but we also must recognize that religious ideology played a role in initiating the tragedy. Like many things, religion is complex and multifaceted, with many different effects. Religion's complexity challenges psychologists to understand its nuances and to communicate that knowledge to our fellow human beings. The attack of September 11, 2001, reminds us of the importance of this task.

REFERENCES

Harris, J. F. (2001, Sept. 14). God gave U.S. 'what we deserve,' Falwell says. *Washington Post*, Page C3.

James, W. (1902/1985). *The varieties of religious experience*. New York: Penguin.

Pargament, K. (1997). *The psychology of religious coping*. New York: Guilford.

Solomon, S., Greenberg, J., & Pyszczynski, T. (2000). Pride and prejudice: Fear of death and social behavior. *Current Directions in Psychological Science, 9*, 200-204.

Washington File, US Department of State International Information Programs (1999). FBI websites document evidence against bin Laden. Retrieved 9/11/01 from: http://usinfo.state.gov/topical/pol/terror/99129502.htm

---

Editor's note: The author is an expert on the psychology of religion. He has a website devoted to this topic at www.psywww.com/psyrelig/fundamental.html.

Effects of Indoor Foliage Plants on Subjects' Recovery from Mental Fatigue

Seiji Shibata
Bunkyo Women's University, Saitama, Japan

Naoto Suzuki
Doshisha University, Kyoto, Japan

We investigated the effects of foliage plants on participants' task performance, fatigue, and mood. Two room conditions (one in which plants were arranged in the room and the other without plants) were created. Undergraduate students (M=33, F=37) performed 2 sessions of a key response task under one of the two room conditions. As for task performance, Plant × Session interaction was significant ($p< .05$). The task scores in the first session did not show any significant difference between plant and no plant conditions. Although it was not significant, the scores in the second session showed higher scores under the plant condition than in the no-plant condition. The plant condition and the no-plant condition did not show any differential effects on the deterioration of task scores in each session. Though the plants affected task scores, they did not show any effects on subjects' moods or fatigue. It was concluded that the presence of the plant might have influenced recovery from mental fatigue.

It is generally believed that looking at "nature" has a positive influence on recovery from fatigue, or stress. For many years this belief was unaccompanied by scientific proof. However, in recent years, the study of possible relations between the human and the natural environment gained popularity, and many researchers have provided scientific data in support of this belief.

Ulrich (1984) reported that patients who could see nature through windows recovered faster than those who could see only a brick building. Hartig, Mang, & Evans (1991) suggested positive benefits on mental fatigue from just a brief exposure to the natural environment. Moreover, Hartig, et al. (1996) showed that people evaluated their mood more positively when they saw slides of natural scenes than those of urban settings. This suggests that the effects of nature can also be conveyed through slides.

Author info: Correspondence should be sent to: Seiji Shibata, Department of Human Studies, Bunkyo Women's University, 1196 Kamekubo, Oimachi, Iruma-gun, Saitama-ken 356-8533, Japan; or to sshibata@hum.u-bunkyo.ac.jp.
North American Journal of Psychology, Vol. 3, No. 2, 385-396
© NAJP

There are two mainstreams of thought about the effects of nature on recovery from fatigue. The first was offered by Kaplan (Kaplan, S., 1995; Kaplan, S. & Kaplan, R. 1989), who theorized that the scenes of nature awakens one's fascination with it: clouds, sunsets, snow patterns, the motion of the leaves in the breeze — these readily capture one's attention, but in an undramatic fashion. The presence of natural beauty enables one to rest.

The second (Ulrich, 1981, 1984; Ulrich et al., 1991) maintains that recovery from fatigue through the natural environment is one of several patterns of spontaneous emotional responses elicited by visual information such as moderate depth, moderate complexity, the presence of a focal point, and the presence of contents such as vegetation or water. Perception of these properties prompts a shift toward more positive emotional states and blocks negative emotions and thoughts.

People can be exposed to natural elements such as trees and plants in various ways, not only in the forest or the wilderness, but also indoors. Kaplan (1993) investigated the role of nature in the workplace environment. She found that when people could see trees through the window, it was beneficial for them even where many of those trees were small, but that no psychological benefit was found when people could see only buildings through the window, even when natural lighting came through the window. In indoor settings, the window provides the main contact with nature. However, when the scenery from the window is not enough, people try to compensate for it. For example, they might do this by decorating the room with posters of natural scenes (Heerwagen & Orians, 1986). In recent years, it has become popular to add foliage plants to improve the amenity of the indoor environment. This can also be a reflection of people's needs of natural elements as Heerwagen & Orians (1986) suggested.

In the workplace, it has become common to arrange plants in the hope that the arrangement would reduce the amount of such symptoms as eye-strain, shoulder stiffness, back pain and brain fatigue, or that it would help one recover from these symptoms. Kondo & Toriyama (1989) found that seeing plants would help subjects recover from the visual fatigue caused by the constant use of a visual display terminal (VDT). Moreover, Asaumi, Nishina, Nakamura Masui, and Hashimoto (1995) found that seeing plants during the VDT task helped reduce visual fatigue, and seeing plants after the task helped the participants recover from visual fatigue. These studies used critical flicker fusion frequency (CFF) as an index of fatigue and the main focus of the study was on visual fatigue. Therefore, there was no way to show how the performance of the VDT task had changed due to the presence or absence of plants.

The relationship between task performance and the density of plants in the work setting was studied by Larsen, Adams, Deal, Kweon, and Tyler (1998). They found that as plant density increased task performance grew worse, but the mood of the workers grew better. Superficially, their results contradicted expectations, but their results could have been influenced by the type of task. The task they used was a simple repetitive task to identify the letters in a document. Stone & Irvine (1993, 1994) investigated the effects of having access to a window on task performance, and found that the influence of windows on one's performance depends on the nature of the task. They found that performance on simple repetitive tasks such as filing or computing was made worse by the view from the window, while the windowed environment enhanced performance on creative tasks. It was also possible that the existence of plants distracted participants' attention from the task in Larsen et al.'s (1998) study.

In the everyday work setting, it is common to take a short rest during work. There was no break in the Larsen et al. study. Therefore, it is not known how plants affected the recovery of task performance after the break. Since VDT tasks in the office require gazing at the display, the existence of the plants may distract one from doing the task. However, we expected that the presence of the plants would have greater influence on the workers during break time because they would pay more attention to their surroundings then. That is, we predicted that the presence of the plant would have a greater effect on recovering from fatigue than on task performance. In the current study, therefore, we used a key-typing task with a VDT, and placed an interval between the repeated tasks to investigate the effects of the plants on the participants' fatigue and recovery from fatigue.

METHOD

Participants

Seventy undergraduate students (male 33, female 37), who volunteered by signing a roster, participated in this study. The students received extra credit for their participation.

Task Environment

The room used in this study was 210cm × 360cm, and had one door and no window. A 180cm × 180cm screen was placed 150cm away from the door to divide the room into two parts. The control devices were arranged on the door side, and the rest of the room was used for a working room. In the work place, there were two shelves of 91cm(W) × 62cm(H) × 28.5cm(D) each, a desk, and a chair. One of the shelves was placed at the corner most distant from the door (northwestern corner), facing south. The other shelf was placed at the corner of the opposite side (northeastern corner), facing west. The desk was 100cm(W) × 64.5cm(H) × 69cm(D), and was placed against the east wall on the longer side of the desk and 100cm away from the north wall. A 17-inch computer display and a keypad were placed on the desk, and two speakers were on both sides of the display. There were two conditions, one with the plants arranged in the room, and the other in which there were no plants. The number of plants was not too dense to distract the subjects' attention in doing their task but dense enough for identifying the influences of the presence of the plants. Potted plants used in this study were of three different kinds. They were heart leaf philodendron (Philodendron scandens oxycardium), augusta (Strelitzia-augusta), and walking stick palm (Phycorapis Singaporensis). The plants ranged in height from about 15cm to 30cm. The augusta and the philodendron were arranged on the shelf in the northwestern corner, and the stick palm on the shelf in the northeastern corner. These plants were arranged so as not to be in the participants' sight while doing their task.

Task

A key response task was used in this study. Three green lights were shown on the middle of the monitor, and each light was given a number, 1 to 3, below each light. The left light was given the number 1, the middle one was given a number 2, and the right one was given the number 3. The size of each light was 2cm in diameter. The lights were placed on the position that the top of the each light was 6.5cm from the top of the monitor, and the interval between each light was 2.5cm. The center of the middle light was centered horizontally. One of the three lights was turned on at a time, and the task was to push the key corresponding to the number of the lighted one as quickly and as accurately as possible. The duration of the lighting was 200msec, and the interval between two lights was 500msec. When the light was turned on, a 400Hz sine wave was presented for 200msec from the speaker to warn subjects that the light was turned on. The display of the stimuli and the recording of the subjects' response were all controlled by a personal computer in the control room.

Questionnaire

Fourteen items of 10cm analogue scale was created to assess the subjects' mood and fatigue. The items used were "relaxed," "fatigued," "vigorous," "tense," "active," "clear-minded," "happy," "eye-strained," "confident," "sodden," "at ease," "sleepy," "concentrated," and "distracted". In addition, ten items of 7-point SD scales were created to assess the evaluation of the room. The items used were "relaxing -

restless," "comfortable - uncomfortable," "noisy - silent," "spacious - small," "concentrating - distracting," "suitable for a task - unsuitable for a task," "ease - unease," "unpleasant - pleasant," "monotonous - diverse," and "natural - artificial."

Procedure

Participants were assigned randomly to the plant or the no-plant condition. The number of participants in the plant condition was 35 (16 male and 19 female), and the number in the no-plant condition was 35 (17 male and 18 female).

The experiment consisted of two task sessions, each of which lasted 5 minutes with a 3-minute break between tasks. After participants entered the room, they were asked to take a seat and were given task instructions followed by a one-minute practice session with their task. After that, they were asked to place a short vertical line along each of the 14 lines to indicate how they currently felt about their mood and fatigue. When the participants completed the questionnaire, they engaged in the key response task for five minutes. Five minutes later, the subjects were asked to complete the mood and fatigue questionnaire again. They had a 3-minute break, and were asked to rest and instructed not to stand up while they rested. The experimenter was out of the room during the rest. When the time was over, the experimenter returned and the participants were again asked to complete the mood and fatigue questionnaire. Following this, they engaged in the key response task for another five minutes. Again, participants were asked to complete the mood and fatigue questionnaire, and then asked to evaluate the room used in the experiment.

RESULTS

Task performances

Task scores were calculated as follows: we started with correct response rates during the first two minutes (session 1) and those from the last two minutes (session 2) of the first task, those of the first two minutes (session 3) and those from the last two minutes (session 4) of the second task, and all of these were then subjected to the inverse sine transformation. Data from participants whose task scores were more than three standard deviations below the mean were excluded from data analysis, leaving data from 68 participants in the final data set.

Using a 2 (plant) × 4(sessions) × 2(gender) analysis of variance, we found the plant-by-session interaction significant ($F(_{3,192})=2.98$, $p<.05$) and main effect of session significant ($F_{3,192}=47.93$, $p<.01$). Further

TABLE 1 Factor pattern of 14 items of the questionnaire about subjects' mood and fatigue after varimax rotation.

	Factor Loading		
Variable	Factor1	Factor2	Factor3
Active	.809		
Happy	.754		
Clear-minded	.715		
Vigorous	.696		
Confident	.552		
Sleepy		.746	
Sodden		.727	
Eye-strained		.650	
Fatigued		.570	
Distracted		.557	
Relaxed			.753
Tense			-.719
At ease			.711
Concentrated			

analysis showed no plant effects in each of the session points. When the contrasts against the first session point were examined, however, there were significant effects in session 3 and session 4 ($F_{1,64}=7.04$, $p<.01$; $F_{1,64}=5.54$, $p<.05$, respectively). As illustrated in Figure 1, difference scores of the second task (session 3 and session 4) were higher in the plant condition than in the no-plant condition.

TABLE 2 Transition of the Factor Scores of Each Three Factors

		Session			
Factor		1	2	3	4
Energy	M	0.261	-0.139	-0.035	-0.087
	SD	(0.849)	(0.914)	(0.850)	(0.934)
Fatigue	M	-0.137	0.064	-0.118	0.191
	SD	(1.009)	(0.845)	(0.785)	(0.786)
Relax	M	0.385	-0.660	0.472	-0.197
	SD	(0.818)	(0.782)	(0.590)	(0.737)

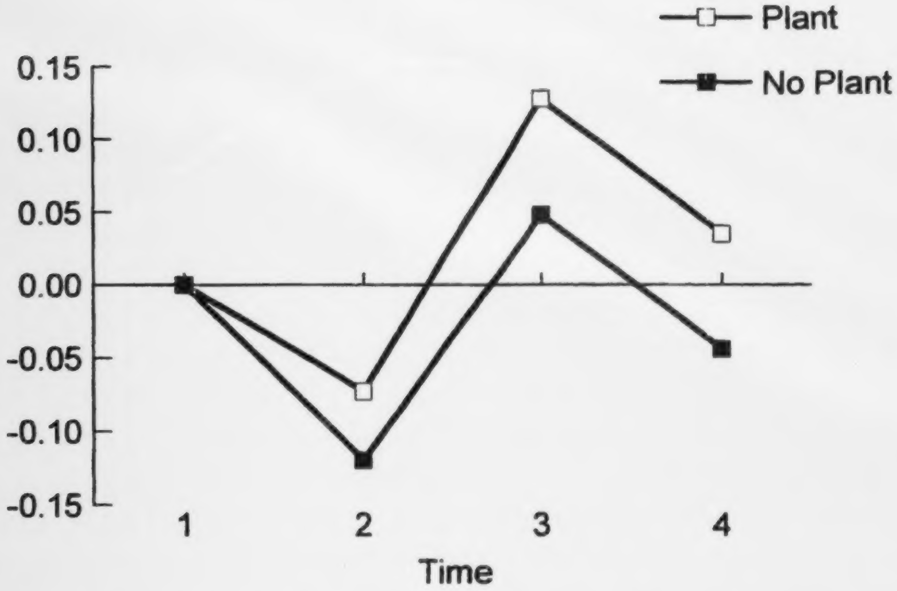


FIGURE 1 Transition of the Task Score with the Score at Time 1 as the Baseline.

Mood and fatigue questionnaire

Score on the mood and fatigue questionnaire was defined as the length from the left end of the line to the marker placed by each participant for each item. Factor analysis with a varimax rotation was carried out with the scores thus defined on 14 mood and fatigue items. To facilitate interpretation, factor loadings greater than or equal to 0.5 are shown in Table. Three factors emerged. For Factor 1, those items with high loadings were "vigorous", "active", "clear-minded", "happy", and "confident", which seems to suggest the "energy factor" as a possible name. For Factor 2, the high loadings were on "fatigued", "eye-strained", "sodden", "sleepy", and "distracted", which could be called the "fatigue factor". For Factor 3, the high loadings were on "relaxed," "tense," and "at ease," which could be named the "relax factor." Table 2 shows mean factor scores at each point of time. Each set of factor scores was subjected to the 2(plant) × 2(gender) × 4(session) analysis of variance. The main effect of session was significant in all three factors ($F_{3,192}$ =5.80, $p<.01$ for the energy factor; $F_{3,192}$ =3.97, $p<.05$ for the fatigue factor; $F_{3,192}$ =49.97, $p<.01$ for the relax factor). However, no main effects of Plant or Gender were found. Scores on the energy factor decreased significantly during the first session of the task ($F_{1,64}$ =9.71, $p<.01$ for the change from session 1 to session 2), and scores of the fatigue factor increased significantly during the second portion of the task ($F_{1,64}=15.96$, $p<.01$ for the change from session 3 to session 4). The scores on the relax factor changed significantly between all session points, decreased and increased alternately ($F_{1,64}=70.53$, $p<.01$ from session 1 to session 2; $F_{1,64}=120.90$, $p<.01$ from session 2 to session 3; $F_{1,64}=43.27$, $p<.01$ from session 3 to session 4).

Room Evaluation

The room evaluation scores showed a significant main effect of the presence of the plant ($F_{10,56}=2.29$, $p<.05$) in the Plant × Gender multivariate analysis of variance (MANOVA). The mean evaluation item scores were shown in Figure 2. The stepwise discriminant analysis showed that the main effect of plant presence was significant on the evaluation scores of "noisy - silent" and "spacious - small" ($F_{1,56}=4.89$, $p<.05$; $F_{1,56}=4.51$ $p<.05$, respectively). That is, the room with the plants was evaluated as quieter and smaller compared to the room without the plants.

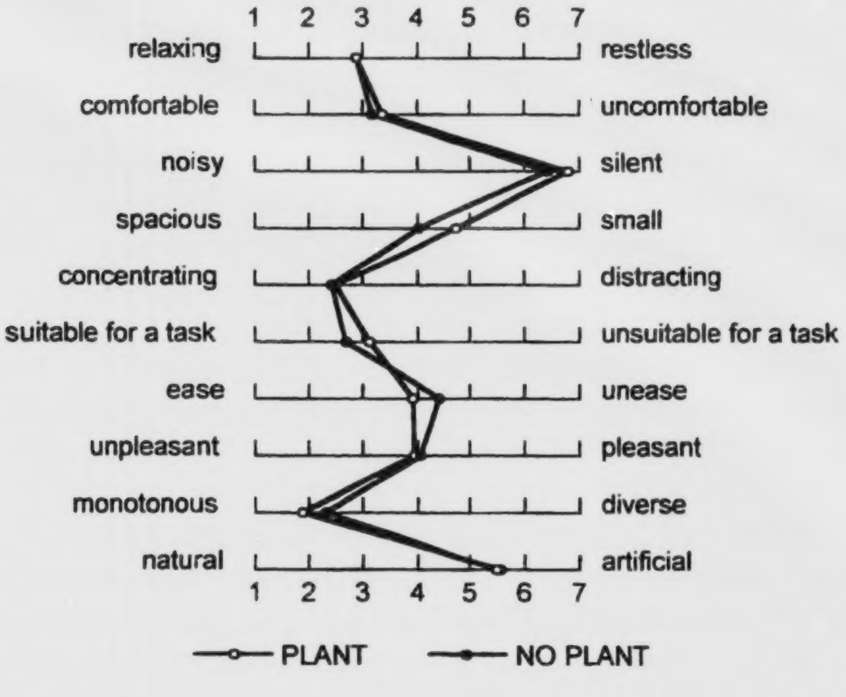


FIGURE 2 Profiles of the Evaluation of the Room in Each Condition

DISCUSSION

In this study, effects of the presence of the plant were seen on the evaluation of the room and task performances. However, there was no significant effect of the plants on the mood or fatigue evaluation. In the mood and fatigue evaluation, the score on the energy factor decreased first, and then the score on the fatigue factor increased. This seems to suggest that the fatigue might have accumulated gradually over the period of the experiment. The room was evaluated as being smaller under the plant condition than in the no-plant condition. Perhaps visual information was perceived as denser in the plant condition than the no-plant condition, leading to the perception of the room as smaller in the former than in the latter. Also, the evaluation of the room being more "silent" in the plant condition than in the no-plant condition suggests that attention to the surrounding noise decreased because the subjects' attention was drawn to the plants in the room.

Task performance effects of the plants were observed in the plant condition in the second task. The fact that the scores on the second task were higher in the plant condition than in the no-plant condition appears to suggest that the recovery from fatigue could have been accelerated by the plants. We should also note that the plants did not affect scores on the first task or the second task. This seems to suggest that the existence of the plants affected the restoration from fatigue rather than the reduction of fatigue from the VDT task. Kaplan (Kaplan, S., 1995; Kaplan, S. & Kaplan, R. 1989) suggested that the unintentional shift of attention is useful for the restoration from fatigue, and natural beauty attracts people's attention unintentionally. The potted plants in the room might have awakened pleasant feelings in participants and might have helped reduce their mental fatigue as Kaplan says in his attention restoration theory. In the current study, plants were arranged so as not to interfere with task performance. Thus it is possible that the plants could have helped suppress fatigue. However, as the results of Larsen et al. (1998) and Stone & Irvine (1993, 1994) suggest, exposure to the information irrelevant to the task is likely to distract participants' attention to the task, and hinder their task performance rather than reduce their fatigue.

Ulrich (Ulrich, 1981, 1984; Ulrich et al., 1991) suggested that recovery from fatigue is one of the emotional and aesthetic response patterns to nature, and that recovery happens because the positive mood induced by nature reduces the stress. However, the current study showed restorative effects by plants on task performance, without showing any effects on mood evaluation. The results of Tennessen and Cimprich's (1995) study showed that the restoration of directed attention was not affected by mood, and Kaplan pointed out that the restoration of directed attention is different in this respect from stress reduction, the goal of which is to improve mood. Taking these studies into consideration, it seems proper to interpret our results to mean that the restoration from fatigue was due to restoration of attention, not stress reduction.

This study investigated the effects of the presence of plants on participants' mood and task performance in the work setting. Results showed that plants had a restorative effect on fatigue in task performance, but that there were no effects of plants on mood and fatigue evaluation. Although this is speculative, it appears as though participants' attention was restored by the existence of plants, considering that the task used in this study was very demanding. It is possible that even a few plants in the work setting may help subjects in recovering from fatigue caused by a VDT task.

REFERENCES

Asaumi, H., Nishina, H., Nakamura, H., Masui, Y., & Hashimoto, Y. (1995). Effect of ornamental foliage plants on visual fatigue caused by visual display terminal operation. *Journal of Shita, 7*, 138-143.

Asaumi, H., Nishina, H., Namba, R., Masui, Y., & Hashimoto, Y. (1995). Evaluation of impression of ornamental foliage plants and psychological rating of rooms with ornamental foliage plants by means of semantic differential method. *Journal of Shita, 7*, 34-45.

Hartig, T., Böök, A., Garvill, J., Olsson, T. & Gärling, T. (1996). Environmental influences on psychological restoration. *Scandinavian Journal of Psychology, 37*, 378-393.

Hartig, T., Mang. M., & Evans, G. W. (1991). Restorative effects of natural environment experiences. *Environment and Behavior, 23*, 3-26.

Heerwagen, J. H. & Orians, G. H. (1986). Adaptations to windowlessness: a study of the use of visual decor in windowed and windowless offices. *Environment and Behavior, 18*, 623-639.

Kaplan, R. (1993). The role of nature in the context of the workplace. *Landscape and Urban Planning, 26*, 193-201.

Kaplan, R. & Kaplan, S. (1989). *The experience of nature: A psychological perspective.* New York: Cambridge University Press.

Kaplan, S. (1995). The restorative benefits of nature: toward an integrative framework. *Journal of Environmental Psychology, 15*, 169-182.

Kondo, M., & Toriyama, T. (1989). Experimental research on the effectiveness of using green in reducing of visual fatigue caused by VDT operation. *Journal of the Japanese Institute of Landscape Architecture, 52*, 139-144.

Larsen, L., Adams, J., Deal, B., Kweon, B-S., & Tyler, E. (1998). Plants in the workplace: The effects of plant density on productivity, attitudes, and perceptions. *Environment and Behavior, 30*, 261-281.

Stone, N. J., & Irvine, J. M. (1992). Performance, mood, satisfaction, and task type in various work environments: A preliminary study. *The Journal of General Psychology, 120*, 489-497.

Stone, N. J., & Irvine, J. M. (1994). Direct or indirect window access, task type, and performance. *Journal of Environmental Psychology, 14*, 57-63.

Tennessen, C. M.. & Cimprich, B. (1995). Views to nature: effects on attention. *Journal of Environmental Psychology, 15*, 77-85.

Ulrich, R. S. (1981). Natural versus urban scenes: some psychological well-being. *Environment and Behavior, 4,* 17-23.

Ulrich, R. S. (1984). View through a window may influence recovery from surgery. *Science, 224*, 420-421.

Ulrich, R. S., Simons, R. F., Losito, B. D., Fiorito, E., Miles, M. A., & Zelson, M. (1991). Stress rocovery during exposure to natural and urban environments. Journal of Environmental Psychology, *11*, 201-230.

Acknowledgements: We would like to acknowledge the useful comments from Professor Shizuhiko Nishisato, University of Toronto.

Adolescent Substance Abuse Behavior: An Empirical Examination of the Covalent Security Attribution Model of User and Non-User Groups.

David Daniel Bogumil
Wright State University

The attribution of the causality regarding the perceived quality of community and safety differs between user and non-user groups of illegal drugs. This study is an evaluation of a proposed model of Covalent Security Attribution. An evaluation of the attribution theoretical model and intergroup relations of users and non-users of illegal drugs reveals differences in group cognitive consistency of their attribution of "locus of control." The concepts referenced herein as "locus" and "control" of casual perceptions affects the perceptions of safety and quality of community life. A confirmatory factor analysis of the factorial invariance of the Covalent Security Attribution measurement model demonstrates configural and metric invariance. An analysis of the structural model using structural equation modeling indicates that the Covalent Security Attribution model reveals an important linkage between cognitive function of group consistency and reciprocity of intergroup perception.

The linkage between causality and control in the attribution process posits a "naive scientist" (Heider, 1958), an "intuitive scientist" (Kelley, 1967), and/or an "applied scientist" (Kelley, 1972) seeking an alignment of internal states with external phenomena (Rotter, 1966). Heider (1944, 1958) provides the seminal framework for the arguments of attribution as a process of evaluation of psychological states in a dimensional space.

Weiner's rational attribution taxonomy as applied to group attribution in intergroup research within organizational behavior research (Martinko, 1995; McDonald, 1995; Weiner, 1995) explains the attribution process as a function of interdependent group social identity (Brewer, 1985; Brewer & Schneider, 1990; Deschamps, 1973-1974; Doise et al., 1988; Tajfel, 1972, 1973) within an organizational environment.

The external factors of guns and weapons, widespread drug use, and the dearth of crime within the neighborhood environment and the internal factors of helplessness, lack of control, and anomie both directly affect the individual's perception of security and safety. The members of the

---

Author info: Correspondence should be sent to: David Daniel Bogumil, Ph.D., Dept. of Sociology and Anthropology, Wright State University, Millett Hall, 3640 Col. Glenn Highway, Dayton, OH 45435 david.bogumil@wright.edu
North American Journal of Psychology, 2001, Vol. 3, No. 3, 397-414.
© NAJP

BEST COPY AVAILABLE

user and the non-user groups maintain their schema by rational attribution of causality of their status. Furthermore, their status reifies the casual schema that informs their rational attributions regarding their group, the other group, and the environment (Branscombe, Schmitt, & Harvey, 1999; Oetting, 1999; Weiner, 1998; Cialdini, 1996). Users and non-users coexist; however, their perception of security in their shared enviroment is differentially affected by their user status--a user status based on experience in their community's drug culture.

Weiner (1979, p. 4) states that the attribution process includes an individual's accumulation of experiences that are tantamount to subconscious inferences that are directly linked with self-esteem and self-concept. Rosenberg (1989) evaluated the component of self-esteem for adolescents and concluded that contextual factors play a significant role in the management and development of self-esteem. Branscombe et al. (1999), using Rosenberg's scale items, found a negative relationship between ingroup positive attribution towards an outgroup and ingroup self-esteem. Branscombe et al. claim the context of the ingroup idenification was found to serve as a factor in the attribution process.

This study's hypotheses evaluate the measurement and structural models that are derived from the conceptual model (see Conceptual Model section) and are presented following the Measures section of this study.

METHOD

Participants

The Youth Evaluation Study is a study of drug use of students from the Consolidated Metropolitan Statistical Area (CMSA) of Chicago--Cook County, Illinois and Gary--Lake County, Indiana. This study's sample of 1,021 students was comprised of the 5th grade students who are the graduating class of the year 2000. The representation of diversity of racial and ethnic characteristics of the population in the sample was accomplished by stratification of schools based on state enrollment records that include students' demographic characteristics. The sampling frame resulted in 31 elementary, 8 secondary (middle), and 8 high schools inclusion in this study.

All students completed the questionnaires in-class and follow-up questionnaires were administered to absentee students. The questionnaires were available in English and Spanish language versions. The questionnaire required approximately fifty minutes to complete.

Questionnaire items regarding frequency of drug use were adapted from other national drug use surveys. The drug use project Monitoring the Future provided many items included in this study (Bachman, Johnston, & O'Malley, 1991; Johnston, O'Malley, & Bachman, 1993). The coding method used in the instrumentation recorded frequency of drug use study (Bachman et al., 1991; Johnston et al., 1993). A participant was classified as a drug user when usage occurred during a six month period prior to the students' response. The non-user group ethnicity: 32.8% White, 29.9% Black, and 32.9% Latino. The user group ethnicity: 32.6% White, 23.9% Black, and 38.6% Latino. Gender composition for each group is statistically equivalent: 54% female, 46% male (additional demographics--see Bogumil, 1999).

Conceptual Model

The conceptual model presented is a three-factor model of attribution and security. I will call this modeled process the Covalent Security Attribution (CSA) model. This conceptual model emerges from the nexus of Weiner's (1986) attribution theoretical framework as well as the effort of other scholars (see Bogumil, 2001).

This examination of the Covalent Security Attribution model proposes an evaluation of some of the pivotal dimensions of Weiner's model coupled with the related work of Rotter (1966), Heider (1958), and Heiss (1981) used to develop the structure of the CSA model.

This study presents an evaluation of Covalent Security Attribution model limited to the study of perceptions of groups. The study of group perception in a cross cultural and reciprocal environment is a distinct theoretical venture and empirical choice of a "proving ground" (Hui & Triandis, 1985). This study's context or empirical "proving ground" permits an examination of the interaction effects of two groups coexisting in the same environment and/or context. Furthermore, each group defines and reifies the other group's perception of security.

The Covalent Security Attribution Model offers a cognitive "deep structure" model based social roles (Heiss, 1981) that is well suited to the tenets of structural equation modeling. Hofer, Horn, and Eber (1997, p. 249) assert the more parsimonious structure of a model (one with less parameters) provides for more viable and critical test of theoretical proposition (Bentler & Mooijaart, 1989). Moreover, the CSA model presents an empirical test of the identification of the external and internal states that affect the perceptions of the two groups.

User and non-users groups that occupy the same environment and function within a shared social context of security. Hayduk (1996, p. 3) contends that "only one identifiable position from the literature" should be tested in a structural model—a structural model devoid of the burden of "complex composites of competing conceptual views." Non users and users of illegal drugs perceive their shared environment based on their perceptions of the other. Their perceptions of their shared environment are products of both their external and internal attributions. These attribution states manifest different perceptions of security of their shared environment. This CSA model posits that only attribution states determine the perception of security.

Measures

The indicator manifest variables are all Likert scaled variables. Rosenberg (1989) evaluated the component of self-esteem for adolescents (Bachman et al., 1991; Johnston et al., 1993; Lewis, Volk, Edwards, & Schulenberg, 1996). The *Rosenberg Self-Esteem Scale* items parallel the indicators of the Internal Stable Uncontrollable latent variable. The indicator variables for the latent variable Internal Stable Uncontrollable (ISU) are: ICR, "I feel I cannot do anything right;" ING, "Sometimes I think I am no good at all;" and ILU, "I feel that my life is not very useful."

The indicator variables for the latent variable External Stable Uncontrollable are: GWMN, "Guns and other weapons often seen in my neighborhood;" MPND, "Many people in my neighborhood use illegal drugs;" and TLCN, "There is a lot of crime in my neighborhood."

A group's shared value of the utility of continuity of life events is measured by the their attitudes and global judgments regarding the safety and quality of their neighborhood. The indicator variables for the latent variable SAFE are: NGPL, "Neighborhood is a good place to live" and IFSN, "I feel safe in my neighborhood."

User and non-user groups were formed by evaluating student reports of illegal drug usage. Illegal drug use includes the use of inhalants, marijuana, heroin, cocaine, crack, and methamphetamine. The coding method used in this study's instrumentation recorded the frequency of drug use (Bachman et al., 1991; Johnston et al., 1993; Lewis, Volk, Edwards, & Schulenberg, 1996). Drug use was coded as positive when usage occurred during six months prior to the students' response. A dichotomous variable of drug usage was established to identify the user and non-user groups. The user group (n=256) was comprised of students who had reported the use of illegal drugs during the past six months. The non-user group (n=765) was comprised of students who reported no use of illegal drugs during the past six months

Measurement Model

The measurement model posits a factorial invariant model for both user and non-users. The conceptual model asserts structural differences between user and non-users. The exogenous latent variables of Internal

Stable and Uncontrollable (ISU), External Stable and Uncontrollable (ISU), and the dependent latent variable SAFE were evaluated by a confirmatory factor analysis using AMOS 4.0 (Arbuckle, 1994-1999; Keith et al., 1995; Keith, 1997; MacCallum, 1995). The assessment of the measurement model was accomplished by conducting a confirmatory factor analysis of the three-factor model. Meredith (1993), among others, concluded the establishment of factorial invariance using two or more groups from a population selected without representative sampling does not reduce the rank of the matrix providing selection of the two groups was not based on the dependent variable (Hofer et al., 1997, p. 250).

Factorial invariance may be evaluated at several levels (Meredith, 1993; Cheung & Rensvold, 1999; Rensvold & Cheung, 1998; Hui & Triandis, 1985). The various levels of invariance are progressively more restrictive in the constraints placed upon the model under evaluation. The incremental steps in the levels of assessment of factorial invariance include configural invariance (models that are identical in arrangement of parameters with individual assessment of each group), metric invariance (factor loadings), invariance of item intercepts (item bias or differential item functioning), and invariance of residual (error) variances.

The goal of full factorial invariance between two groups requires that one succeeds in obtaining fit indices indicative of a good-fit of the groups (Byrne, 2001; Mulaik, James, Van Alstine, & Bennett, 1989). An unconstrained model is compared to a baseline model with a ratcheting of constraints on the two-group (User and Non-user) model until an *a priori* level of factorial invariance is obtained. The desiderata of full factorial invariance of measurement model for two groups has been suggested as the optimum environment for evaluation of the structural model. However, partial invariance of covariance models has been submitted as a reasonable standard for the evaluation of multiple group models (Aitken, Stein, & Bentler, 1994; Byrne, 1994; Meredith, 1993; Rensvold & Cheung, 1998, p. 1022).

The evaluation of user and non-user groups "metric invariance" is established by the assessment of the goodness of fit of the model to the data with the factor loadings for both groups invariant. The manifest variables or indicator variables λs of the construct or latent variable ξ must employ the same metric across groups as a basis for evaluation of causal relationships among the constructs or latent variables ξs. The evaluation of factorial invariance of the two-group model of Covalent Security Attribution will examine the following hypotheses.

1) The two-group measurement model demonstrates configural invariance.

2) The two-group measurement model demonstrates metric invariance.

The conceptual model of the structural relationship among the latent variables for each model supported the hypothesized differences in structural relationships of the user and non-user groups. The Covalent Security Attribution structural model further examines the additional four hypotheses.

3) The user group's attribution represented by their perception of their Internal, Stable, Uncontrollable (ISU) state will be a positive function of their perception of security (SAFE).
4) The user group's attribution represented by their perception of their External, Stable, Uncontrollable (ESU) state will be a neutral function of their perception of security (SAFE).
5) The non-user group's attribution represented by their perception of their Internal, Stable, Uncontrollable (ISU) state will be a negative function of their perception of security (SAFE).
6) The non-user group's attribution represented by their perception of their Internal, Stable, Uncontrollable (ESU) state will be a negative function of their perception of security (SAFE).

RESULTS

The analysis of the proposed Covalent Security Attribution (CSA) model involved a confirmatory factor analysis of the measurement model. The confirmatory factor (i.e., the *a priori* hypothesized measurement model) analysis was performed on the total sample of students (*n*= 1021). The results of the confirmatory factor analysis are presented in Figure 1. The correlation matrix and standard deviations are presented in Table 1.

The results of the structural model for the non-user group are presented in Figure 2. The results of the structural model for the user group are presented in Figure 3. The examination of the structural models indicated that the four hypotheses were confirmed. The method of estimation of the covariance matrix was maximum likelihood. Multiple fit indices are presented to capture various aspects of model viability and performance. All parameter estimates were significant. The model identification process requires that selected moments (manifest variables) be set to unity (Byrne, 2001).

Figure 1:
CFA Three Factor Measurement Model of Attribution States
(n=530)

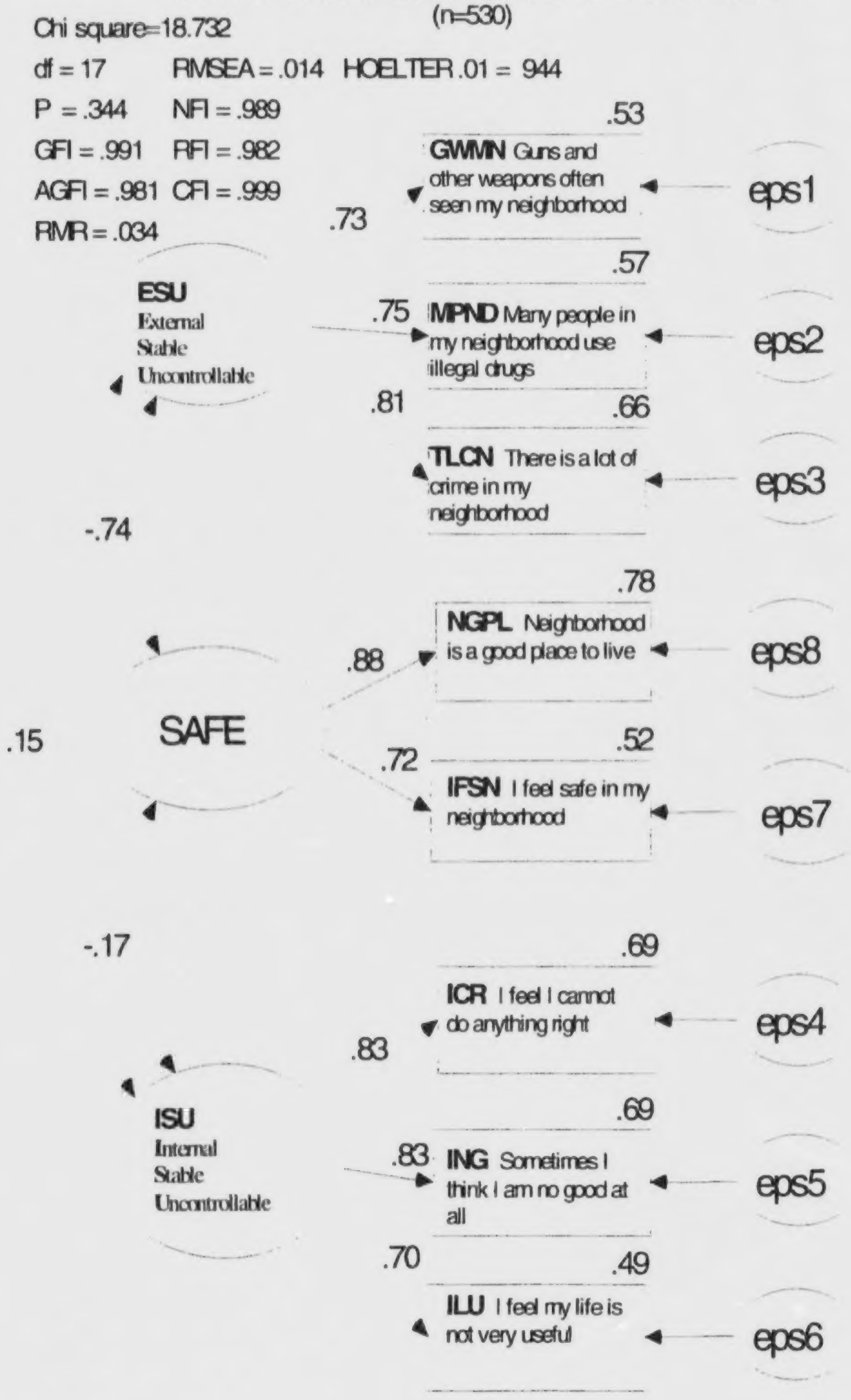

Chi square=18.732
df = 17
RMSEA = .014
HOELTER .01 = 944
P = .344
NFI = .989
GFI = .991
RFI = .982
AGFI = .981
CFI = .999
RMR = .034
.53
GWMN Guns and other weapons often seen my neighborhood
eps1
.73
ESU
External
Stable
Uncontrollable
.57
.75
MPND Many people in my neighborhood use illegal drugs
eps2
.81
.66
TLCN There is a lot of crime in my neighborhood
eps3
-.74
.78
NGPL Neighborhood is a good place to live
eps8
.88
SAFE
.15
.52
.72
IFSN I feel safe in my neighborhood
eps7
-.17
.69
ICR I feel I cannot do anything right
eps4
.83
ISU
Internal
Stable
Uncontrollable
.69
.83
ING Sometimes I think I am no good at all
eps5
.70
.49
ILU I feel my life is not very useful
eps6

Figure 2:
Structural Model for Student Non-Users
(n=765)

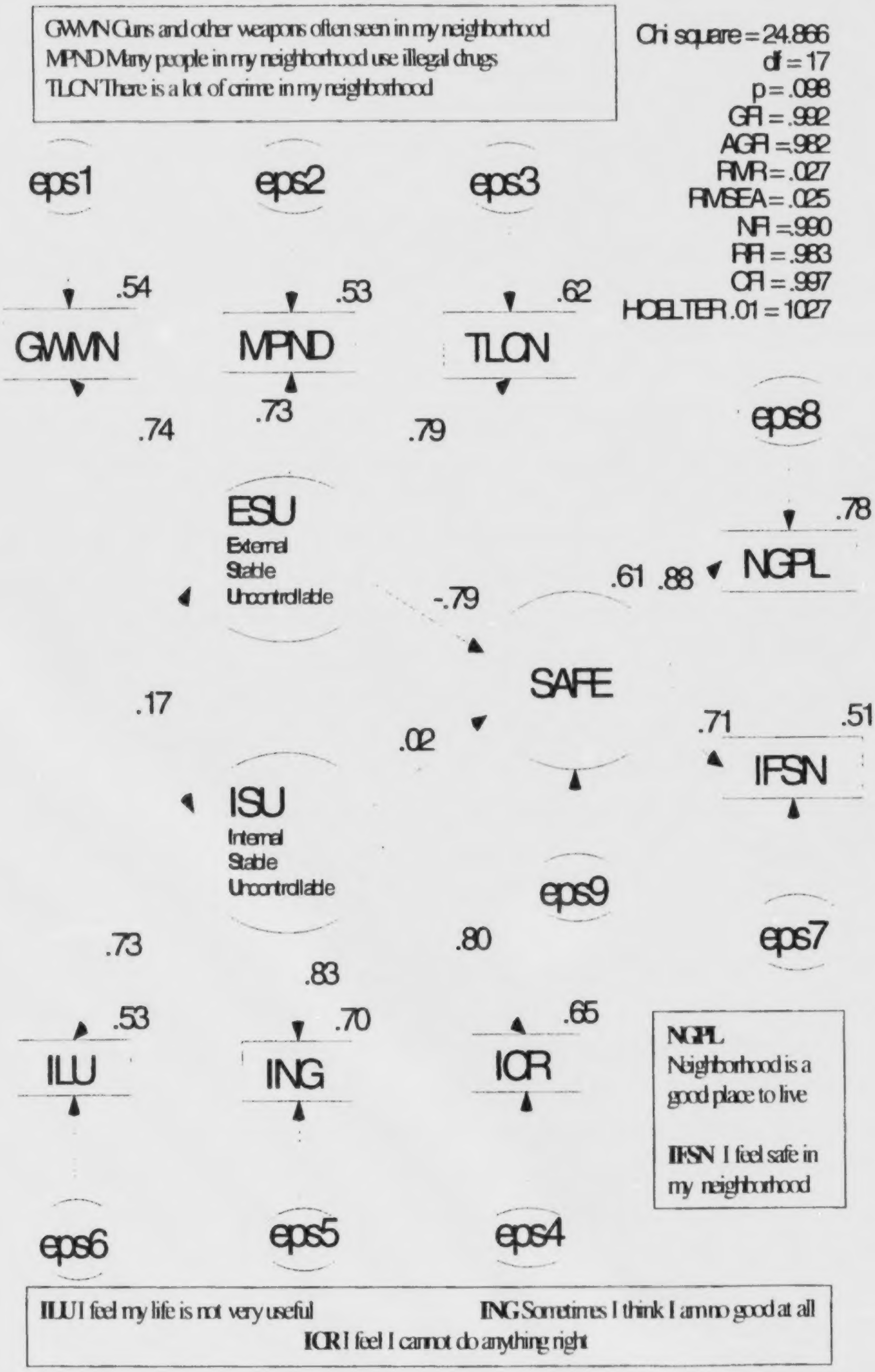
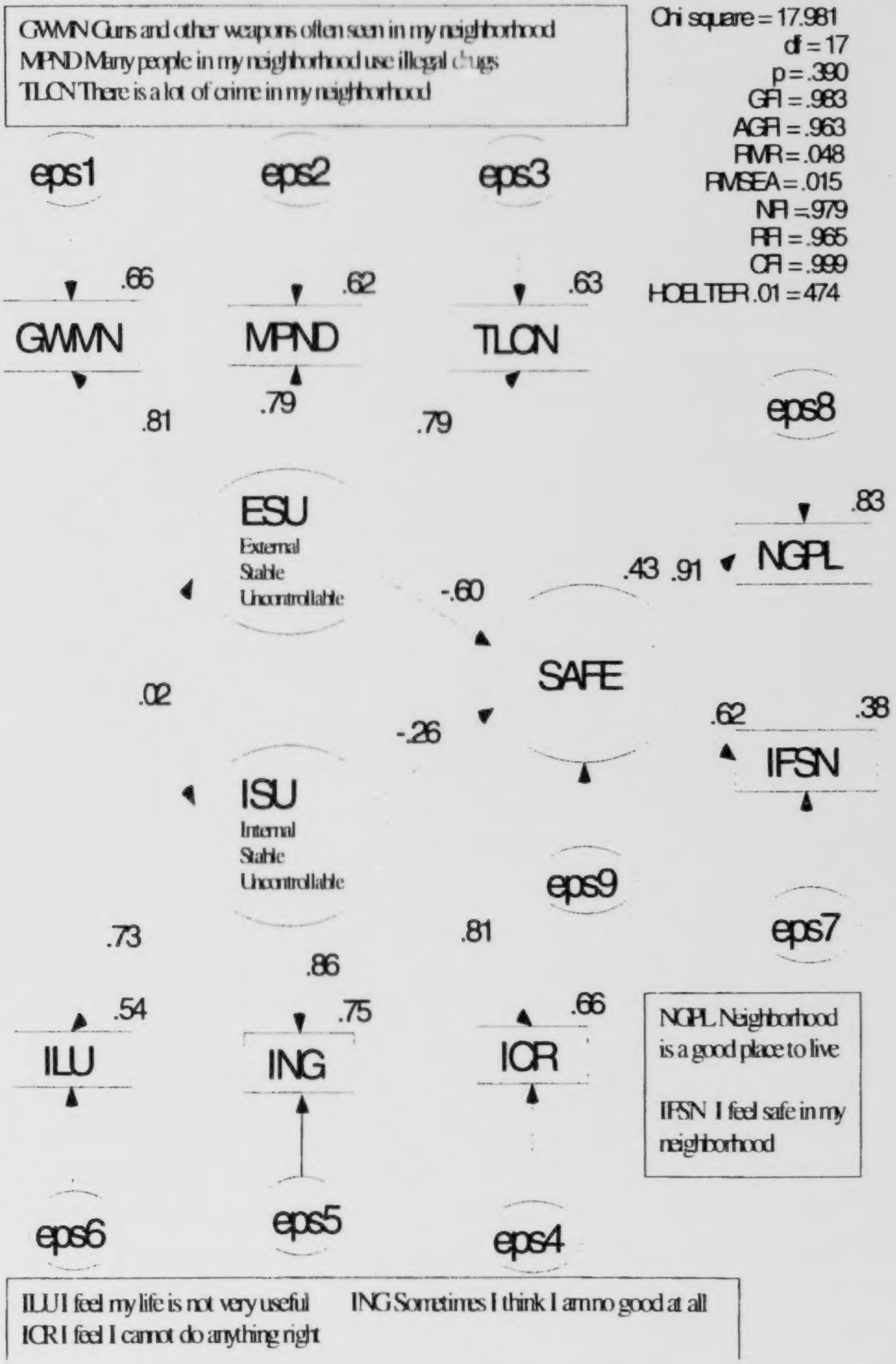

GWMN Guns and other weapons often seen in my neighborhood
MPND Many people in my neighborhood use illegal drugs
TLCN There is a lot of crime in my neighborhood
Chi square = 24.866
df = 17
p = .098
GFI = .992
AGFI = .982
RMR = .027
RMSEA = .025
NFI = .990
RFI = .983
CFI = .997
HOELTER .01 = 1027
eps1
eps2
eps3
.54
.53
.62
GWMN
MPND
TLCN
.74
.73
.79
eps8
.78
ESU
External
Stable
Uncontrollable
NGPL
-.79
.61
.88
SAFE
.17
.02
.71
.51
IFSN
ISU
Internal
Stable
Uncontrollable
eps9
eps7
.73
.80
.83
.53
.70
.65
ILU
ING
ICR
NGPL
Neighborhood is a good place to live
IFSN I feel safe in my neighborhood
eps6
eps5
eps4
ILU I feel my life is not very useful
ING Sometimes I think I am no good at all
ICR I feel I cannot do anything right

Figure 3:
Structural Model for Student Users
(n=256)



GWMN Guns and other weapons often seen in my neighborhood
MPND Many people in my neighborhood use illegal drugs
TLCN There is a lot of crime in my neighborhood

Chi square = 17.981
df = 17
p = .390
GFI = .983
AGFI = .963
RMR = .048
RMSEA = .015
NFI = .979
RFI = .965
CFI = .999
HOELTER .01 = 474

NGPL Neighborhood is a good place to live

IFSN I feel safe in my neighborhood

ILU I feel my life is not very useful
ING Sometimes I think I am no good at all
ICR I feel I cannot do anything right

eps1
eps2
eps3
.66
.62
.63
GWMN
MPND
TLCN
.81
.79
.79
eps8
.83
ESU
External
Stable
Uncontrollable
.43
.91
NGPL
-.60
SAFE
.02
-.26
.62
.38
IFSN
ISU
Internal
Stable
Uncontrollable
eps9
eps7
.73
.81
.86
.54
.75
.66
ILU
ING
ICR
eps6
eps5
eps4

The confirmatory analysis suggested a good fit of the measurement model (see Figure 1). The user and non-user groups were evaluated in a multiple groups comparison to test the "configural" invariance of two-group measurement model (Byrne, 2001; Byrne et al, 1989). The two-group unconstrained measurement model produced a good fit of data.

TABLE 1 Correlation Matrix and SDs

Pearson Correlation Coefficients

	NGPL	GWMN	IFSN	TLC	MPND	ICR	ING	ILU
NGPL	1							
GWMN	-.51 ***	1						
IFSN	.61 ***	-.37 ***	1					
TLC	-.52 ***	.59 ***	-.42 ***	1				
MPND	-.48 ***	.59 ***	-.35 ***	.60 ***	1			
ICR	-.13 ***	.11 ***	-.15 ***	.09 **	.09	1		
ING	-.09 **	.08	-.14 ***	.09 **	.11 ***	.69 ***	1	
ILU	-.11 ***	.10	-.16 ***	.12 ***	.12 ***	.60 ***	.63 ***	1
SD	1.26	1.44	1.34	1.34	1.43	.83	.91	.88

*** P < 0.001 level (2-tailed).

** P < 0.01 level (2-tailed).

N = 1, 021 The .01 significance level was used because of the large number of correlations.

The two-group model's factor loadings were then constrained to equality to test the "metric" invariance of two-group measurement model (Byrne, 2001; Byrne et al., 1989). Bollen (1989, p. 366) and Cheung and Rensvold (1999, p. 2) suggest equivalence or invariance of the factor loadings must be established before proceeding to the further evaluation of a model. The factor loadings for the item responses (manifest variables) for the same constructs (latent variables) across both groups were evaluated for equivalence. The hypothesis of factorial invariance of the measurement model between the user and non-user group was confirmed (see Figure 1).

DISCUSSION

The groups within this study have experienced significant intergroup dynamics. Additionally, the social influence of these groups within a reciprocal social environment of school and community has reified the social identity of each group (Deschamps, 1973-1974; Doise, Turner, Rabbie, & Horwitz, 1988; Greene, 1999; Lang, 1987; MacCallum, 1995; McDonald, 1995; Tajfel, 1972, 1973; Vourlekis, 1999; Young, 1987). Moreover, the internal group dynamic's of each groups's members incubates and manifests the group's social identity. The evaluation of group identity is the composite of several levels of social characteristics such as: sex, age, ethnicity, and religious beliefs (Brower & Garvin, 1989).

The salient social characteristics that are shared among individual group members are, in this study, also inculcated in the attribution states (internal-external, stable-unstable, & controllable-uncontrollable) of group members (see citations to Wiener and Wiener and associates in literature review section). This study finds that the user group member's perception of their group identity is a manifestation of the internal user group dynamics—the user group's dynamics foster attribution states that formulate, create, and reify the user group's identity.

Notably, the reification of user group identity is also a product of the interaction with the non-user group. The groups are functioning within the interdependent larger society. Notwithstanding, there are people who use drugs and people who do not use drugs—these groups are mutually exclusive and exhaustive.

The compulsory environment of the school and the shared environment of community are not optional or mutually exclusive settings. These settings are the forced existence of every school-aged adolescent. Adolescents cannot reasonably find sanctuary and live a separate reality apart from the community, the school, or the other group. This study's findings suggest that drug use status is extremely significant in the process of group identification and cohesion. Furthermore, the internal attribution states of members of either group are the critical lens that reveal the intragroup affects of our shared drug culture.

The user group member's internal states, social identifiable characteristics, and drug use status reveal a portrait of group dynamics based in hopelessness of a social identity that was chosen through drug usage. A user group identity is sustained by the non-user group's indifference, fear, loathing, scapegoating, and desire to preserve their own identity as pure and without culpability for the user group's continued existence and behavior.

The non-user group members' internal states, social identifiable characteristics, and drug use status reveals a portrait of group dynamics based in the hope of a social identity that was chosen through abstention from drug usage and sustained by the user group's volatile presence in the school and community. The user group member's perceive the non-user group members with indifference, fear, and loathing. Subsequently, the non-user group's presence preserves the user group's identity as a corrupted form of purity—a corrupt purity that places one in touch with the harsh reality of life and not some inexperienced view of hope—as expressed by non-users.

Scholars (Brower & Garvin, 1989; Greene, 1999; Lang, 1987; Vourlekis, 1999; Young, 1987) have suggested that variations in "individual attributes" (e.g., "age, sex, cognitions, affects, or previous experiences") of group member can have significant effects on member interaction (Brower & Garvin, 1989, p. 93-94). Moreover, this study's evaluation of the attributional states (internal-external, stable-unstable, & controllable-uncontrollable) of group members may require a design that permits the control for these "relevant individual attributes" and "group conditions" (Brower & Garvin, 1989, p. 93-94). Thereby, the attributional states of group members may be examined without the expectations of undue influence from these other variables.

This study's sample reveals that numerous social demographic variables provide a basis of nearly equally matched characteristics or "relevant individual attributes" between both "natural" groups and within "natural" groups (Brower & Garvin, 1989). An examination of sex, age, ethnicity (Black, Latino, & White), co-residence of both mother and father, educational objectives, achievement effort, cognitive activities, physical activities, and participation in family responsibilities, indicates an equal distribution within both user and non-user groups and between user and non-user groups. Therefore, the analysis of the attributional states as presented in this study embody the exposition of the intragroup specific interactions.

The perception of external stable and uncontrollable attributions of the non-user group indicates a strong and negative effect on the non-user group's perceptions of security (see Figure 2). However, the internal, stable and uncontrollable perception of non-user attributions of self indicate a neutral effect on the non-user group's perception of security. This result alone would suggest the non-user group's state of helplessness in their perception of environmental security. Furthermore, the non-user's perception of the external factors affecting their perception of security are strongly negative. However, the context of this shared environment with users of illegal drugs suggests that the social influence (i.e., threat) of a smaller group of users exceeds the social influence of the larger non-user group.

The perception of external stable and uncontrollable attributions of the user group indicates a strong and negative effect on the user group's perceptions of security (see Figure 3). The perception of a lack of security by the user group coupled with the user group's activities that both implicitly and explicitly contribute to their group's perception of insecurity is ironic. The user group's perception of security is moderately strong compared to that of the non-user group. This difference between the non-user and user group's perception of security represents the user group's increased knowledge and awareness of the activities that contribute to less security in the environment and the user group's potential to avoid, evade, and manage such a threat to their security. Nonetheless, the non-user group's perception of lack of security within their shared environment seems authenticated by the user group's perception and direct awareness of external factors.

The user group's perception of self is most disturbing. This group's perception of self has a moderate and negative effect on their perception of security. Moreover, the user group demonstrates that the higher their groups' self esteem the greater potential for a lack of security. The brinkmanship of conflict and the high potential for a spiral of negative reciprocity of action mainifests a perception that affects the user group's security. The user group perceives user group self esteem as a potential threat to the security of their environment, whereas the non-users perceive their self esteem as unimportant to their security.

There are additional intragroup dynamics that manifest perceptions of consensus and cohesion within each group. The non-user group's variations in their collective perception of esteem, control, and security are artifacts of intragroup dynamics processes of consensus and cohesion. The matched social and demographic characteristics of each group, with the requisite exception of substance use status, presents a refined, pragmatic, and utilitarian narrowing of the scope conditions pertinent to the evaluation of the intragroup dynamics—intragroup attribution states. Furthermore, the central proposition of this study, the Covalent Security Attribution (CSA) model, is illuminated by revealing the fundamental connections between the intragroup attribution states, the intragroup elements of consensus and cohesion, and intragroup dimensions: esteem, control, and security.

The non-user groups' intragroup perception of esteem is based in their shared behavior of abstention from drug use. Moreover, their capacity to resist the culture of drug use present in their community and schools authenticates their capacity to control their behavior. The non-user group's consensus actions create a sense of cohesion that results in an intragroup perception of security. However, this group does not exist in isolation for the user group. Therefore, the non-user group's perceptions and attributions are tempered by their constant interactions with the user group.

The user group's intragroup perception of esteem is based in their shared behavior of drug use. Their capacity to resist the social and legal sanctions and prohibition present in their community and schools authenticates their capacity to control their behavior. The user group's consensus actions create a sense of cohesion that results in an intragroup perception of security—the delusional perception of the intragroup invulnerability of deviance. Again, this group does not exist in isolation from the non-user group. Therefore, the user group's perceptions and attributions are tempered by their constant interactions with the non-user group.

The internal dynamics of the group by definition will always be identified by the characteristics of behavior or demographics that set the group apart and manifest its social identity as a group. Clearly, non-users and users are the only two categories of individuals available for group membership within our drug culture. They are exclusive groups. Therefore, their drug use status and perpetuation of pathological attribution states within the drug culture of school and community fosters an interdependent and reciprocal status as an artifact of their intragroup and intergroup dynamics. This negative reciprocity and cycle of the degeneration of humanity may be symptomatic of the remains of a "society [that] has reinforced the more narcissistic and pleasure seeking aspects of human striving rather than the more altruistic and socially useful" (Garvin, 1984, p. 18).

Lastly, the Covalent Security Attribution model has significant value to the literature in group behavior (Aitken et al., 1994; Bavelas, Hastorf, Gross, & Kite, 1965; Branscombe el al., 1999; Brewer, 1985; Brewer & Schneider, 1990; Brown & Wooten-Millward, 1993; Kashmima & Kashmina, 1993), international negotiation (Bogumil, 1993; Druckman & Broome, 1991; Druckman & Harris, 1990; Patchen & Bogumil, 1995; 1997; Stoll & McAndrews, 1986), and communities with substance abuse populations (Bachman et al., 1991; Donohew et al., 1999; Johnston et al., 1993; Miller & Volk, 1996; Oetting, 1999; Volk, Edwards, Lewis, & Schulenburg, 1996).

REFERENCES

Aitken, A. C., Stein, J. A., & Bentler, P. M. (1994). Structural equation analyses of clinical subpopulation differences and comparative treatment outcomes: Characterizing the daily lives of drug addicts. *Journal of Consulting and Clinical Psychology, 62*, 488-499.

Arbuckle, J. L. (1994-1999). Amos (Version 4.0) [Computer software]. Chicago: SmallWaters Corporation.

Bachman, J. G., Johnston, L. D., & O'Malley. P. M. (1991). *Monitoring the Future project after seventeen years: Design and procedures* (Monitoring the Future Occasional Paper No. 33). Ann Arbor: MI: Institute for Social Research.

Bavelas, A., Hastorf, A. H., Gross, A. E., & Kite, W. R. (1965). Experiments on the alterations of group structure. *Journal of Experimental and Social Psychology, 1*, 55-70.

Bentler, P. M. & Mooijaart, A. (1989). Choice of structural model via parsimony: A ratonale based on precision. *Psychological Bulletin, 106*(2), 315-317.

Bogumil, D. D. (2001). *Attribution states: Emergency room crisis management.* Paper presented at the 96th Annual Meeting of the American Sociological Association. Regular Session: Health and Well Being. Anaheim, CA.

Bogumil, D. D. (1993). Models of conflict and cooperation: Interaction patterns of international actors. *Dissertation Abstracts International, 55*(03A), 745. (University Microfilms No. 9420786). Doctoral dissertation. Purdue University.

Bollen, K. A. (1989). *Structural equations with latent variables.* Wiley Series in Probability and Mathematical Statistics. New York: John Wiley and Sons.

Branscombe, N. R., Schmitt, M. T., & Harvey, R. D. (1999). Perceiving pervasive discrimination among African Americans: Implications for group identification and well-being. *Journal of Personality and Social Psychology, 77*(1), 135-149.

Brewer, M. B. (1985). The role of ethnocentrism in intergroup conflict. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations*, (2nd ed., pp. 88-102). Chicago: Nelson-Hall.

Brewer, M. B. & Schneider, S. K. (1990). Social identity and social dilemmas: A double-edged sword. In D. Abrams & M. A. Hogg (Eds.), *Social identity theory: Constructive and critical advances*, (pp. 169-184). New York: Springer-Verlag.

Brower, A. M. & Garvin, C. D. (1989). Design issues in social group research. *Social work with groups, 12*(3), 91-102.

Brown, R. & Wooten-Millward, L. (1993). Perceptions of group homogeneity during group formation and change. *Social Cognition, 11*(1), 126-149.

Byrne, B. M. (2001). *Structural equation modeling with AMOS.* Mahwah, NJ: Lawence Erlbaum Associaties, Publishers.

Byrne, B. M. (1994). Testing for the factorial validity, replication, and invariance of a measurement instrument: A paradigmatic application based on the Maslach Burnout Inventory. *Multivariate Behavioral Research. 29*, 289-311.

Byrne, B. M., Shavelson, R. J. & Muthen, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*(3), 456-466.

Cheung, G. W. & Rensvold, R. B. (1999). Testing factorial invariance across groups: A reconceptualization and proposed new method. *Journal of Management, 25*(1), 1-27.

Cialdini, R. B. (1996). Social influence and the triple tumor structure of organizational dishonesty. In D. M. Messick & A. E. Tenbrunsel (Eds.), *Codes of Conduct: Behavioral research into business ethics* (pp. 44- 58). New York:

Deschamps, J. (1973-1974). Attribution, social categorization and intergroup representations. *Bulletin de Psychologie, 27*(13-14), 710-721.

Doise, W., Turner, J. C., Rabbie, J.M., & Horwitz, M. (1988). Individual and social identities in intergroup relations. *European Journal of Social Psychology, 18*(2), 99-111.

Donohew, L., Clayton, R. R., Skinner, W. F., & Cole, S. (1999). Peer networks and sensation seeking: Some implications for primary socialization theory. *Substance Use & Misuse, 34*(7), 1013-1023.

Druckman, D., & Broome, B. J. (1991). Value differences and conflict resolution: Familiarity or liking. *Journal of Conflict Resolution, 35*(4) 571-593.

Druckman, D., & Harris, R. (1990). Alternative models of responsiveness in international negotiation. *Journal of Conflict Resolution, 34*(2), 234-251.

Garvin, C. D. (1984). The changing contexts of social group work practice: Challenge and opportunity. *Social work with groups, 7*(1), 3-19.

Greene, R. R. (1999). Person-in-environment, human behavior theory, and social work method. In R. R. Greene & P. H. Ephross (Eds.), *Human behavior theory and social work practice* (pp. 1-30). Hawthorne, NY: Aldine De Gruyter.

Hayduk, L. A. (1996). *LISREL issues, debates, and strategies*. Baltimore, MD: The John Hopkins University Press.

Heider, F. (1944). Social perception and phenomenal causality. *Psychological Review, 51*, 358-374.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Heiss, J. (1981). Social roles. In M. Rosenberg & R. H. Turner (Eds.), *Social psychology* (pp. 66-93). New York: Basic Books, Inc.

Hofer, S. M., Horn, J. L., & Eber, H. W. (1997). A robust five-factor structure of the 16PF: Strong evidence from independent rotation and confirmatory factorial invariance procedures. *Personality & Individual Differences, 23*(2), 247-269.

Hui, C. H. & Triandis, H. C. (1985), Measurement in cross-cultural psychology: A review and comparison of strategies. *Journal of cross-cultural psychology, 16*(2), 131-152.

Johnston, L. D., O'Malley, P.M., & Bachman, J. G. (1993). *National results of drug use from the monitoring the future society, 1975-1992, Vol. 1, Secondary students, NIH Publication No. 93-3597.* Rockville, MD: National Institute on Drug Abuse.

Kashmima, E. S. & Kashmina, Y. (1993). Perceptions of general variability of social groups. *Social Cognition, 11*(1), 1-21.

Keith, T. Z. (1997). Using confirmatory factor analysis to aid in understanding the constructs measured by intelligence tests. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 373-402). New York: Guilford.

Keith, T. Z., Fugate, M. H., DeGraff, M., Diamond, C. M., Shadrach, E. A., & Stevens, M. L. (1995). Using multi-sample confirmatory factor analysis to test for construct bias: An example using the K-ABC. *Journal of Psychoeducational Assessment, 13,* 347-364.

Kelley, H. H. (1967). Attribution in theory in social psychology. In D. Levine (Ed.), *Nebraska symposium on motivation, Vol. 15.* Lincoln, NE: University of Nebraska Press.

Kelley, H. H. (1972). Attribution in social interaction. In Jones, E. E., Kanouse, D. E., Kelley, H. H., Nisbett, R. E., Valins, S. & Weiner, B. (Eds.), *Attribution: Perceiving the causes of behavior.* (pp. 1-27). Morristown, NJ: General Learning Press.

Lang, N. (1987). Social work practice in small social forms: Identifying collectivity. In N. C. Lang & J. Sulman (Eds.), *Collectivity in social group work: Concept and Practice* (pp. 7-32). New York: Haworth Press, Inc.

Lewis, R. A., Volk, R. J., Edwards, D. W., & Schulenberg, J. E.(1996). *Youth Evaluation Study: Questionnaire.* Unpublished manuscript, Purdue University, West Lafayette, IN.

MacCallum, R. C. (1995). Model specification: Procedures, strategies, and related issues. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 16-36). Thousand Oaks. CA: Sage.

Martinko, M. J. (1995). *Attribution theory: An organizational perspective.* Delray Beach, FL: St. Lucie Press.

McDonald, D. M. (1995). Fixing blame in n-person attributions: A social identity model for attributional processes in newly formed cross functional groups. In M. J. Martinko (Ed.), *Attribution theory: An organizational perspective* (pp. 273-288). Delray Beach, FL: St. Lucie Press.

Meredith. W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58,* 525-543.

Miller, T. Q., & Volk, R. J. (1996). Weekly marijuana use as a risk factor for initial cocaine use: Results from a six-wave national survey. *Journal of Child and Adolescent Substance Abuse, 5*(4), 55-78.

Mulaik, S. A., James, L. R., Van Alstine, J., & Bennett, N. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin, 105*(3), 430-445.

Oetting, E. R. (1999). Primary socialization theory. Developmental stages, spirituality, government institutions, sensation seeking, and theoretical implications. V. *Substance Abuse and Misuse, 34*(7), 947-982.

Patchen, M., & Bogumil, D. D. (1995). Testing alternative models of reciprocity against interaction during the Cold War. *Conflict Management and Peace Science, 14*(2), 163-195.

Patchen, M., & Bogumil, D. D. (1997). Comparative reciprocity during the Cold War. *Peace and Conflict, 3*(1), 37-58.

Rensvold, R. B. & Cheung, G. W. (1998). Testing measurement models for factorial invariance: A systemic approach. *Educational and Psychological Measurement, 58*(6), 1017-1034.

Rosenberg, M. (1989). *Society and the adolescent self-image: Wesleyan edition.* Middletown, CT: Wesleyan University Press.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs, 80*, 1-28.

Stoll, R. J., & McAndrews, W. (1986). Negotiating strategic arms control, 1969-1979. *Journal of Conflict Resolution, 30*(2), 315-326.

Tajfel, H. (1972). La categorization sociale. In S. Mosovici (Ed.), *Introduction a la psychologie sociale* (Vol. 1). Paris: Larousse.

Tajfel, H. (1973). Social identity and intergroup behavior. *Social Science Information, 13*, 65-93.

Volk, R. J., Edwards, D. W., Lewis, R. A., & Schulenburg, J. (1996). Smoking and preference for brand of cigarette among adolescents. *Journal of Substance Abuse, 8*(3), 347-359.

Vourlekis, B. S. (1999). Cognitive theory for social work practice. In R. R. Greene & P. H. Ephross (Eds.), *Human behavior theory and social work practice* (pp. 173-205). Hawthorne, NY: Aldine De Gruyter.

Weiner, B. (1979). A theory of motivation for some classroom experiences. *Journal of Educational Psychology, 71*, 3-25.

Weiner, B (1980). *Human motivation.* New York: Holt, Rinehart & Winston.

Weiner, B. (1995). Attribution theory in organizational behavior: A relationship of mutual benefit. In M. J. Martinko (Ed.), *Attribution theory: An organizational perspective* (pp. 3-6). Delray Beach, FL: St. Lucie Press.

Weiner, B. (1998). Discovering general laws of social motivation. In J.G. Adair, D. Belanger, & K. L. Dion (Eds.), *Advances in psychological science* (Vol. 1, pp. 93-109). East Sussex, UK: Psychology Press.

Young, R. (1987). Collectivity vs. Group: A framework for assessment. In N. C. Lang & J. Sulman (Eds.), *Collectivity in social group work: Concept and Practice* (pp. 33-43). New York: Haworth Press, Inc.

Note: This article was supported by a grant from the National Institute on Drug Abuse (Grant DA06514).

Conjunctive Forms and Conditional Inference in Questions and Statements

Marc E. Pratarelli
University of Southern Colorado

&
Adam Lawson
Florida Southern College

Utterances that use conjunctives such as *and*, *or*, and *either-or* increase cognitive workload because they require more complex processing than simpler ones. To test this notion, two experiments were carried out to examine the effects these conjunctions have on the process of deriving inferences used to judge congruency. In Experiment 1 participants read 3-4 sentence long vignettes to establish a context after which a test sentence, using one of the above conjunctions, was rapidly judged to be a match or nonmatch. Mean reaction times (RT) and accuracy (RA) were analyzed. Experiment 2 essentially replicated the first with the exception that the test phrases were constructed as interrogatives. Results from both experiments were consistent with each other and demonstrated that *and* was least ambiguous, followed closely by *either-or*. However, the *or* operator produced significant delays in response time as well as reducing accuracy by as much as 90 percent in one condition. Results are interpreted in view of applied concerns where investigations rely on asking complex questions to determine facts.

In language, conjunctions serve the purpose of conveying linkage between two or more conceptual entities or propositions. Cognitively, the reader/listener is forced to evaluate one concept, conjoin it with another in some specified form, and then comprehend the sum or product. The use of conjunctions like the words *and, or, either-or*, etc., are used to convey conditional properties of a sentence or question for which the reader/listener must consider alternative interpretations. For example, Springston and Clark (1973) examined the comprehension times for short pseudoimperative clauses and sentences and concluded that the use of *or* could be conceptually reduced to *and* plus a negative, i.e., "*not (p and q)*" (p.269). These sorts of mental operations are critical in certain applied settings, in particular, when interrogations are performed.

Author info: Correspondence should be sent to: Dr. Marc Pratarelli, Psych. Dept., University of Southern Colorado, 2200,Bonforte Blvd., Pueblo, CO 81001.

North American Journal of Psychology, 2001, Vol. 3, No. 3, 415-428.
© NAJP

In the sentence "Be quiet or I'll shoot you" the reader/listener derives the conditional outcome of being shot only if the individual fails to stop talking. Using Springston and Clark's logic, being shot *and* being quiet do not occur together provided the pseudoimperative instruction conveyed by *or* is met. A similar case can be made when using questions as in the example: "Did you sabotage that laboratory or kill the people in it?;" again, *not (p and q)*. If a suspect who allegedly committed the sabotage of the laboratory, but never intended to kill anyone inside, is asked such a question, the interpretation of the conditional property of the conjunction *or* is semantically equivalent to: "I did not sabotage the laboratory *and* kill the people in it." If *p* equals sabotage, and *q* equals murder, then the above question meets the properties of *not (p and q)*. Assuming the truth is, in fact, that the suspect committed the sabotage resulting in the killing of several innocent people, the improper choice of conjunctive allows the suspect to logically conclude that the question or statement, although true, is incongruent with the original intentions when the crime was committed. Thus, asking such questions introduces a measure of ambiguity that may allow an individual to escape the assignment of guilt. It would similarly introduce noise into any investigation or interrogation to detect their deception and concealed knowledge.

The focus of the present study is to examine the degree of ambiguity in the interpretation of complex statements and questions that use conditionals and conjunctions like those illustrated above. Ambiguity in conditional sentences or questions can be evaluated in many different ways (cf., Chase & Clark, 1972; Clark & Chase, 1972; Geis & Zwicky, 1971; Johnson-Laird, 1969; Lakoff, 1971; Richardson & Ormerod, 1997; Springston & Clark, 1973; Staal, 1968; Taplin, 1971). Geis and Zwicky (1971) examined "invited inferences" using the conditional format of *if-then* statements (p.561). The notion of an invited inference is based on falsely deduced interpretations of sentences they constructed using the if-then operators. In their example: "If you take out the garbage, I'll give you a dollar", one reasonable interpretation is the comparatively similar sentence using negations: "If you don't take out the garbage, I won't give you a dollar." Geis and Zwicky pointed out that the latter invited inference is invalid because the original does not imply it necessarily. In the original example the person could fail to take out the garbage (i.e., failing to meet the first condition), but still receive the dollar, and this is an equally reasonable inference to make. One can argue that the ambiguity is itself dependent on both the context of the situation and the sentence construction.

In the present study the conjunctives *and*, *or*, and *either-or* were manipulated in order to assess comprehension and ambiguity. The rationale for using both questions and statements was based on the application of these results to the construction of conditional questions by interrogators, security, and law enforcement examiners. A major concern in any applied area that uses investigative procedures dependent on asking carefully constructed questions is whether the examinee comprehends the exact or intended meaning of the investigator. The unusual circumstances surrounding investigations of fact and the determination of guilt or innocence are such that individuals with incomplete knowledge of events must use an open and unrestricted context to make inferences in complex questions. Conversely, individuals with a complete knowledge database, e.g., the guilty suspect in the sabotage scenario discussed earlier, use a narrow or restricted context because their knowledge is specific. Ray and Findley's (1984) argument suggests that the latter would be more likely to make invalid inferences. Moreover, because a guilty individual seeking to avoid detection understands the intent and consequences of conditional questions or statements, they are most likely to intentionally make invalid inferences as a means of escaping detection by remaining technically truthful.

The working logic for the present experiments is that conjunctives used to convey conditional properties of questions and statements will vary systematically in their ability to produce invalid inferences. The invalid inferences should be reflected in simple measures like comprehension time and accuracy. The conjunction, *and*, should be the least ambiguous because it expresses the simplest (positive) case of additivity of concepts, e.g., given a prior context, *p and q* is a valid conditional when both components are congruent with the pre-established context. The conjunction *or* should produce the most ambiguity because it does not specify whether the positive or negative condition should apply, e.g., *not (p and q)*. Lastly, the conjunctive set *either-or* should remove any need for making alternative inferences because it makes explicit the conditions that must be met to evaluate the question or statement, e.g., *either p or q, but not both*.

EXPERIMENT 1

The purpose of Experiment 1 was to compare and contrast the effects of using the *and* and the *or* conjunctive operations in declarative sentences. In contrast to some previous research (cf., Ray & Findley, 1984; Richardson & Ormerod,1997; Springston & Clark, 1973) that used comparative performance tasks to examine processing differences between independent conditionals, we chose to use comprehension speed and accuracy because each conditional could be evaluated without dependency on the evaluation or comparison of another conditional. Thus Experiment 1 tested the specific hypotheses that (1) different conjunctive forms will affect comprehension time consistent with the relative degree of ambiguity that each manifests during conditional reasoning, and (2) different conjunctive forms will affect the accuracy of correct forced-choice decisions resulting from the process of conditional reasoning.

Method

Participants. Twenty two college undergraduates participated during development and norming of the stimuli. Afterward, 40 test participants (Mean age = 21.9 years; *SD* = 1.1 years) were similarly solicited. Participants were offered extra course credit for their participation. They were screened for handedness (in order to counterbalance match vs. nonmatch response hands), normal or corrected-to-normal vision, normal neurological history, no history of Learning Disabilities, and English as their first and dominant language. In addition, because the test stimuli were normed on a college student population, the test participants were screened for no-less-than-normal comprehension skills for their college age using the Comprehension subtest of the 1981 version of the Wechsler Adult Intelligence Scale-Revised.

Materials. A total of 36 trials divided equally into six different conditions were developed for testing along with six additional trials used as a practice set. Each trial consisted of a short vignette composed of 3-4 short declarative sentences. The average length of the sentences was 8.1 words. Each vignette was paired with a declarative conditional phrase using either the *and*, *or*, or the *either-or* conjunctive operators. The purpose of each vignette was to establish a context against which the conditional phrase could be tested for congruence, i.e., match or nonmatch. Each vignette depicted a simple scenario in which one or more actors were engaged in one or two events. Appendix 1 lists six such vignettes and a corresponding conditional sentence, one for each of the six possible trial type conditions. Thus there were six trials using *and* that were true with respect to their vignette, i.e., *(p and q)*. Six other trials using *and* were false, i.e., *p and (not-q)*. Note that although the false *and* condition could have been constructed using the sentence structure *not-p (and q)*, such a condition would have inflated the reaction times because the violation of the conditional would have occurred at the beginning. Six trials using *or* were true, i.e., *not (p and q)*, and six were false, i.e., *(p and q)*. Similarly, six trials used the *either-or* operator in the match condition; *(not-p) and q*, *p and (not-q)*, and six others were nonmatches divided equally between *p and q*, and *(not-p) and (not-q)*.

Procedure. The six trial-type conditions were randomly ordered in the 36-trial stimulus list. Participants were given six practice trials to become familiar with the task demands, one practice trial for each of the six conditions, and the practice set could be repeated if necessary. During the test session, participants worked with the Experimenter to ensure comprehension of each vignette prior to the presentation of the test phrase. Each trial began with the subject reading the vignette aloud when it appeared on a computer screen. The vignette remained on the screen as long as the subject needed to continue reviewing it, but never more than 30 seconds. Afterward, the screen cleared and the Experimenter required a verbal command from the subject before initiating presentation of the test phrase. The test phrase, requiring a forced-choice (match vs. nonmatch) manual response on a keyboard, was presented at the center of the screen for five seconds during which the subject was required to respond as quickly and as accurately as possible. Computation of the Means for each subject's data employed an individualized analysis window that rejected trials greater than 1.5 SD from their respective Mean.

Results

A fully repeated-measures analysis of variance (ANOVA) using a 3 x 2 design (type of conjunctive x match/nonmatch response type) was applied separately to reaction time (RT) and response accuracy (RA) data. Greenhouse-Geisser corrected F values are reported for all analyses having greater than one degree of freedom. Using RT, there were significant main effects for both independent variables, $F_{(2,78)} = 23.1$, $p < .001$, and $F_{(1,39)} = 9.35$, $p < .007$, respectively. There was also a significant conjunctive by response type interaction, $F_{(2,78)} = 6.04$, $p < .002$. Figure 1a illustrates the interaction effect for which nonmatch trials were always responded to slower than match trials. Moreover, responses to sentences using the *or* conjunction operator were slower for both matching and nonmatching trials. Planned comparisons reflected that indeed the only differences occurred between *or* and the remaining two conjunctives ($p < .05$). There was no difference between *and* and *either-or*. In the *or* conditions, the dispersion was greater as reflected in the doubling of the standard error.

For the accuracy data, the same basic pattern emerged with significant main effects for both independent variables, $F_{(2,38)} = 28.45$, $p < .0005$ and $F_{(1,39)} = 17.88$, $p < .004$, respectively. There was also a significant interaction effect, $F_{(2,78)} = 20.26$, $p < .001$. Figure 1b illustrates the interaction in which match trials were responded to more accurately than nonmatch trials in all three conditions. Planned comparisons also revealed that the principal effects were between *or* and the other two conjunctive conditions ($p < .05$). There was no difference between *and* and *either-or* response categories. An error analysis was not conducted since any trial could have been excluded because (1) it fell outside the 1.5 SD analysis window, (2) the keypress fell outside the

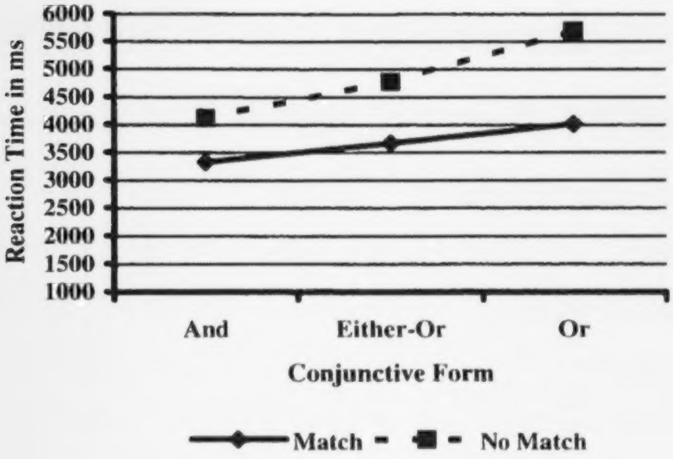


Figure 1a. Reaction times for the different conjunctive forms when the sentence and vignette were congruent (solid line) and incongruent (dashed line).

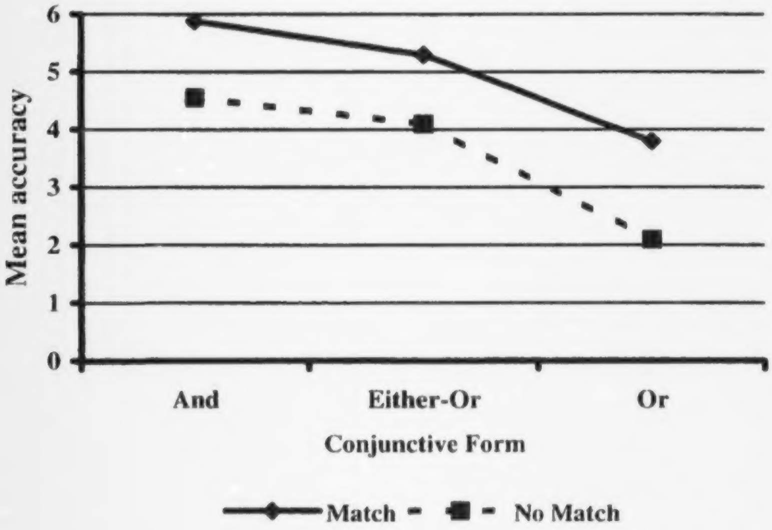


Figure 1b. Mean accuracy out of a possible six correct for the three conjunctive operators. The match condition when sentence and vignette were congruent is designated by the solid line, and the dashed line represents the incongruent condition.

response capture window of eight seconds, and (3) an error occurred in the keypress. Response accuracy across all 40 participants was 76.7 percent, the majority attributable to the *or* condition.

Discussion

These results demonstrate that across these three conjunctive operators, the simple additive case using *and* is easier to comprehend than the *or* case, which Geis and Zwicky (1971) and Springston and Clark (1973) argued was mentally represented as *not (p and q)*. The condition with *either-or* reestablishes the clarity in the inferences drawn from the statements resulting in comprehension times and accuracies that are comparable to the *and* condition. Thus, *p or q* is represented as a negative conditional, i.e., *not (p and q)*, for which Wright and Hull (1986) have suggested people must first conceptualize the positive *and* condition followed by its negation. This can be summarized as a two-step process. The present data suggest that this additional operation requires 1200 to 1300 ms to achieve. In addition, because the accuracy on the *or* trials was significantly poorer, the performance degrades by as much as 50 percent relative to the optimum *p and q* case. Geis and Zwicky (1971) and Springston and Clark (1973) have both argued that the conjunctive *or* activates all the biconditional versions before the correct one can be selected, and this process requires more time and draws on limited processing resources.

The situation for *either-or* is somewhat different, however, because its first component (either) establishes for the listener/reader that there will be two concepts to deal with. Secondly, for the correct inference or conclusion to be drawn, at least one of those two conceptual elements must be true. Two factors appear to be important in the case of *either-or*. The first is that with regard to the issue that at least one of the two elements must be true, participants were not affected when both elements of the argument (*p and q*) were valid. Because our pilot testing showed that using *p but not q*, *not p but q*, and *both p and q* were not substantively different from each other, all three were used interchangeably in the match condition without appearing to draw any additional processing resources. Thus, *either-or* can be said to be the most flexible operator because it permits multiple inferences to be drawn as a function of the three possible conditionals. Given the present results, comparing the *p and q* condition, *either p or q* can be shown to be equivalent to the formula *either: (p and not-q) or (not-p and q) or (p and q).* This latter formula requires that any one of three possible contextual codes, established in working memory by the vignette, needs to exist in order to meet the conditions of *either p or q*.

The final concern to be addressed is the apparent standard offset that occurs between match and nonmatch trial-type responses. Comparing across Figures 1a and 1b, the differential is a relatively fixed amount, about 1000 ms in the RT figure and about 1.5 trials in the RA figure. This is consistent with the notion that mismatches simply take longer to respond to because the evaluation time is longer. There is a comparable performance decrement in accuracy, which suggests these participants were not sacrificing speed for increased accuracy. Instead, both RT and RA registered comparable performance deficits attributable to the increased difficulty of conditional reasoning with incongruent (mismatching) information.

EXPERIMENT 2

The remaining issue to address is whether interrogatives affect the relationship between conditionals. It was also noted earlier that this has important consequences in applied areas where interrogations and legal queries rely on having to ask questions with multiple conceptual elements requiring the use of conjunctives. Therefore, a separate second experiment was conducted to examine differences that might be attributable to the unique case where an individual is questioned as to which alternative is correct.

Method

Participants. A new sample of 40 undergraduate students (Mean age = 20.6 years; SD = .97 years) drawn from the same pool was solicited, screened with the same criteria, and offered extra course credit for their participation.

Materials and Procedures. The same 36 vignettes from Experiment 1 were used with no modifications. The target sentences used for testing the conditionals, however, were altered to form questions with minimal distortion to the existing sentence construction, e.g.,"Chris went jetskiing or hiking with Peter" was changed to "Did Chris go jetskiing or hiking with Peter?" Procedurally, the same protocol used in Experiment 1 was followed.

Results

The same 3 x 2 repeated-measures ANOVA design was applied to these new data. For measures of RT, there were significant main effects for conjunctive type and for response type, $F_{(2,78)} = 18.7$, $p < .0005$ and $F_{(1,39)} = 48.76$, $p < .0005$, respectively. There was also a significant interaction effect, $F_{(2,78)} = p < .002$, illustrated in Figure 2a. Clearly, nonmatch trials were responded to slower than match trials for all three conjunctives. Planned comparisons revealed that the significant pairwise differences occurred between *or* and the two remaining conditions for both match and nonmatch trial-types, and between *and* and *either-or* only in the nonmatch condition. Using measures of RA, there were significant main effects for both independent variables, $F_{(2.78)} = 82.42$, $p < .0005$ and $F_{(2.78)} = 60.18$, $p < .0005$, respectively. There was also a significant

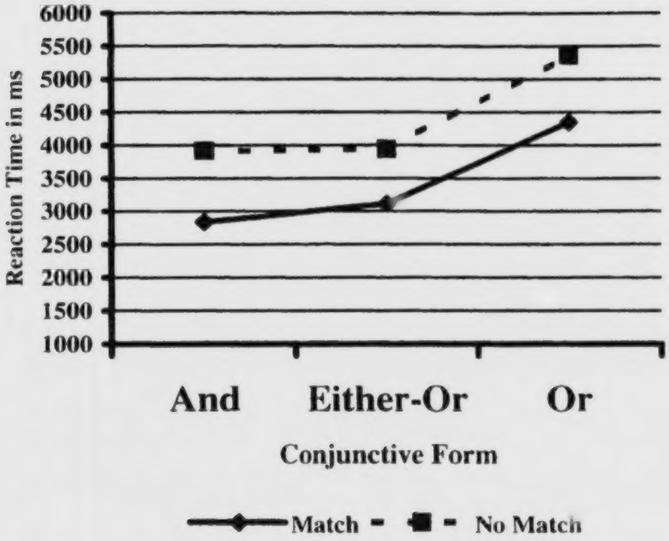


Figure 2a. Reaction times for the different conjunctive forms when the test question and vignette were congruent (solid line) and incongruent (dashed line).

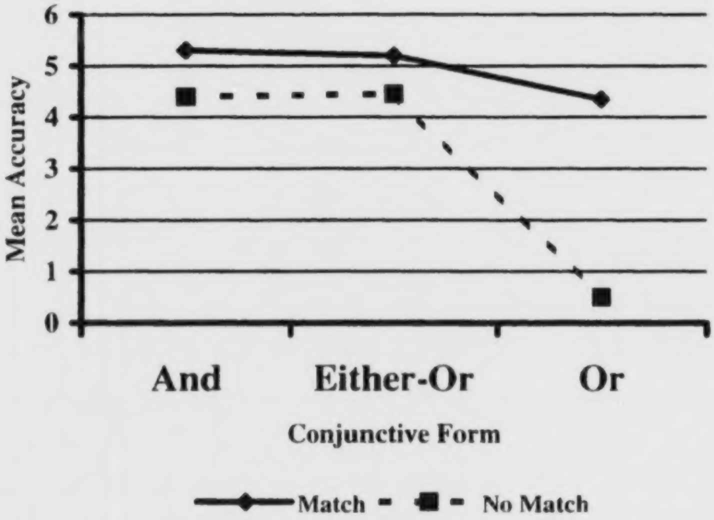


Figure 1b. Mean accuracy out of a possible six correct for the three conjunctive operators. The congruent condition when the sentence and vignette is designated by the solid line, and the dashed line represents the incongruent condition.

interaction effect, $F_{(2,78)} = 32.2$, $p < .0005$, illustrated in Figure 2b. Planned comparisons revealed pairwise differences between *or* and the two other conditions for both match and nonmatch trial-types. In contrast to the RT data, however, there were no differences between *and* and *either-or* in either response-type condition.

Discussion

The focus of Experiment 2 was to examine the effects of different conjunctives on match/nonmatch decisions using interrogatives. These results could then be used to compare and contrast the response patterns elicited by statements (Experiment 1); they will be discussed in the General Discussion section below. For Experiment 2, however, the principal differences are clearly illustrated by Figures 2a and 2b. First, with only one exception, there is again a relatively fixed offset between the match and nonmatch response-type conditions in the RT data. This effect is causally attributable to the delay that results from having to make a nonmatch decision. Consistent with much of the past research, nonmatch decisions are modeled as the process of making a match decision followed by a negation (Clark & Chase, 1972; Just & Carpenter, 1971; Springston & Clark, 1973; Wright & Hull, 1986). Springston and Clark (1973) also noted that negations need not be explicit to have a detrimental impact on processing. The case of *p or q?* is an example of an implicit negation, as compared to the explicit form, *not (p and q)?*. The latter would be a question like "Did you not sabotage the laboratory and kill the people?" By comparison, the *p or q?* question would appear as "Did you sabotage the laboratory or kill the people?" Assuming the individual intended the sabotage, but not the killing, both questions are equivalent. The former has an explicit negation while the latter has an implicit one. In Springston and Clark's model, both samples have *practical* equivalence. That is, they are semantically equivalent to: "I didn't sabotage the laboratory and kill the people."

The upward trend in mean RT seen in Figure 2a reflects the difficulty of assessing the conditional properties of cases other than the basic additive form *p and q?* It also reflects the comparative ease with which *p and q?* is processed. The accuracy data in Figure 2b provides further support for this argument. As difficulty or ambiguity increases, so does the error rate. From the present results, one can conclude that using *or* in a situation that is congruent with the preexisting context would be satisfactory for the intentions of the speaker, assuming the speaker intended *p or q, but not necessarily both*. If the present results were predictive of real-life situations, one could expect a significant but tolerable delay in comprehension, with only a 20 percent risk that the listener will make an incorrect inference. That effect increases radically to 90 percent, however, if one were to use *p or q* incongruently as in "Did you sabotage that laboratory or kill the people in it?" This latter effect is the only notable discrepancy with the results seen in Experiment 1 (compare, e.g., the *or* condition in Figures 1b and 2b). Nevertheless, there is no practical difference between the two. Be they sentences or questions, the incongruent use of *or* will impact comprehension more than any other conjunctive.

GENERAL DISCUSSION

The results from both experiments demonstrate the significant impact of certain conjunctives on conditional reasoning. The inferences that are drawn when using *or* are more likely to be false because of the increased difficulty of selecting the correct representation of its many possible meanings. The *or* operator indeed has considerable flexibility in its usage, but that flexibility comes with a significant cost in lack of specificity. In applied settings where specificity and detail are critical, questions like "Did you sabotage that laboratory or kill the people in it?" should never be used, and *either-or* (e.g., "Did you either sabotage that laboratory or kill the people in it?") should be chosen instead, provided *and* is not appropriate. More importantly, "negatives have almost invariably been found to take longer to process than affirmatives," and this delay may pose problems in situations where it is interpreted as a sign of deception rather than confusion or cautiousness (Springston & Clark, 1973, p.261).

Like the previous research, the present study sought to decompose and isolate the processing deficits of certain conjunctive forms. The results from both experiments are consistent with previous studies despite certain methodological differences needed to accommodate the applied aspects of this research, namely, interrogation and making queries in legal or forensic situations. As is often the case with laboratory research, natural language processing occurs in broader environmental contexts than the sanitized testing suite, and this impacts the validity of the results. Nonetheless, an appreciation for the complexity of the use (and misuse) of conjunctives in natural settings should underscore the importance of the current laboratory results. From subject debriefings, it was clear that the participants in the present study soon focused their attention on the variables being manipulated. However, a routine post-hoc split halves analysis demonstrated there were no differences between the first and latter half of the stimulus set. This suggests that the cognitive-linguistic operations recruited in the process of conditional reasoning where conjunctive forms are used, are highly automatic (Kihlstrom, 1987), and stable. There is little conscious or controlled processing recruited during natural language processing. Therefore, it is reasonable, if not safer, to assume that these laboratory effects generalize to applied settings and should be taken into account when constructing a dialogue or question format during an investigation or discovery of facts.

Future research should seek to fractionate the processing delays and comprehension decrements for the present conjunctives even further, as well as other complex forms like *but*, *unless*, *yet*, and *if*, in order to test thematic or contextual variations and surface structure variations. In addition, the impact on comprehension of conditional reasoning with negations should also be examined further. Given the prevalence with which one hears or reads obtuse surface structures such as "Didn't you know he wasn't there before you, or were you there first?" more effort should be directed at understanding qualitatively the types of comprehension that occur. In the previous example extracted from a videotaped dialogue between two students, the listener appeared to understand because the focus was on the gist of the conversation. Clearly, many others variables, yet to be examined, influence the comprehension of individual linguistic elements.

REFERENCES

Chase, W.G., & Clark, H.H. (1972). Mental operations in the comparison of sentences and pictures. In L.Gregg (Ed.), *Cognition in Learning and Memory.* New York: Wiley.

Clark, H.H., & Chase, W.G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology, 3*, 472-517.

Geis, M., & Zwicky, A.M. (1971). On invited inferences. *Linguistic Inquiry, 2*, 561-566.

Johnson-Laird, P.N. (1969). "&". *Journal of Linguistics, 6*, 111-114.

Just, J.A., & Carpenter, P. (1971). Comprehension of negation and quantification. *Journal of Verbal Learning and Verbal Behavior, 10*, 244-253.

Kihlstrom, J.F. (1987). The cognitive unconscious. *Science, 237*, 1445-1452.

Lakoff, G. (1971). If's and's and but's about conjunction. In C.J. Fillmore, & D.T. Langendoen (Eds). New York: Holt, Rinehart, and Winston.

Ray, J.L., & Findley, P. (1984). Open and restricted context sentences and reasoning with conditional propositions. *Communication Monographs, 51*, 243-252.

Richardson, J., & Ormerod, T.C. (1997). Rephrasing between disjunctives and conditionals: Mental models and the effects of thematic content. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 50A*, 358-385.

Springston, F.J., & Clark, H.H. (1973). *And* and *or*, or the comprehension of pseudoimperatives. *Journal of Verbal Learning and Verbal Behavior, 12*, 258-272.

Staal,, J.F. (1968). "And". *Journal of Linguistics, 4*, 79-81.

Taplin, J.E. (1971). Reasoning with conditional sentences. *Journal of Verbal Learning and Verbal Behavior, 10*, 219-225.

Wright, P., & Hull, A.J. (1986). Answering questions about negative conditions. *Journal of Memory and Language, 25*, 691-709.

Acknowledgements: This research was supported by an institutional research award to the first author. We wish to thank Donald Krapohl for having drawn our attention to this applied concern, and to Pam Cooper, John Griffin, and Jeremy Charles for their assistance in developing the vignettes and collecting the data for the study. We also wish to thank Blaine Browne for constructive criticisms on earlier versions of the manuscript. Direct reprint requests EITHER to Marc E. Pratarelli, Ph.D., Department of Psychology, 2200 Bonforte Blvd., University of Southern Colorado, Pueblo, CO 81001, OR Adam Lawson, Department of Psychology, Florida Southern College, 111 Lake Hollingsworth Drive, Lakeland, FL 33801.... but not both.

APPENDIX 1 Sample Test Sentences and Corresponding Vignettes

1. True (and) ***The man committed autotheft and kidnapping.***
Brenda drove her car and picked up her child. Then she stopped at a gas gas station. Brenda went in to use the restroom. While inside, a man stole her car with her child inside.

2. False (and) ***Karl sold secrets and then told his coach.***
Karl plays football for an NFL team. He doesn't like sitting on the sidelines. He sells the team's secret game plan to an opposing team. When his coach confronted him, he denied selling secrets to the team.

3. True (either-or) Tony bought either a fake ID or a passport.
Tony bought a fake ID from a friend. He used it to get into a bar. Once inside, he drank beer and listened to music. Tony got in trouble with the police.

4. False (either-or) ***Poindexter chases either cats or cars.***
Poindexter the dog likes to chase squirrels up trees at the park. He barks at them for hours sometimes. Then he goes home to his doghouse and falls asleep. He is a happy dog.

5. True (or) ***Chris went jetskiing or hiking with Peter.***
Chris took his friend Peter jetskiing. Peter sat on the back while Chris drove the jetski. When they were done, they both went swimming.

6. False (or) ***Cindy went to the park or broke her toe.***
Cindy went to the water park. While there, she broke her big toe on the water slide. She was taken to the hospital where Doctors taped up her foot.

Monitoring Children's Behaviour in a Remote Community Before and Six Years After the Availability of Broadcast TV

Tony Charlton Ronald Davie
Charlie Panting Mick Abrahams
Cheltenham & Gloucester College of Higher Education

Lilla Yon
Education Department, St Helena, South Atlantic

Teachers rated nursery class children's behaviour 18 months before (i.e. in 1993), and 68 months after the availability of broadcast TV (i.e. in 2000) on the island of St Helena. Across a period of seven years - and almost six years of TV - findings showed few significant differences between teachers' ratings of the two cohorts. In particular, with the 2000 cohort there was no evidence of increases in the types of anti-social behaviours that TV is often alleged to encourage.

Claims are made regularly that TV-viewing encourages anti-social behaviour (e.g. Smith, Nathanson, & Wilson, 1999). However, the evidence underpinning these allegations is frequently misconstrued or flawed. For example, strongest claims for adverse effects emanate from laboratory inquiries. Whilst these studies demonstrate the processes through which children can observe and imitate behaviours viewed on TV, the findings lack ecological validity. Studies tend to measure only immediate or short-term behaviours and to exclude natural consequences of (mis)behaviour. Support for laboratory findings is forthcoming from some field studies (e.g. Boyatzis, Matillo, & Nesbitt, 1995). However, these results need cautious interpretation as participants are often atypical (e.g. youngsters from institutions, and others who are initially more aggressive), thus calling into question the generalization of the results. Elsewhere, outcomes from naturalistic investigations are equivocal (e.g. Williams, 1986; Wiegman, Kuttschreuter, & Baarda, 1992). Moreover, whilst studies of the impact on TV-naïve communities have in some instances revealed short-term behavioural effects, these have either not been sustained over time (Coldevin & Wilson, 1985); have varied among different community members (Granzberg, 1985); or have been inadequately substantiated by weak TV-viewing measures (Williams, 1986).

Author info: Correspondence should be sent to: Dr. Tony Charlton, Cheltenham & Gloucester College of Higher Education, Box 220, Cheltenham, Gloucestershire, GL50 2QF. UK

North American Journal of Psychology, 2001, Vol. 3, No. 3, 429-440.

© NAJP

Furthermore, and most importantly, naturalistic research into TV effects has been hampered by limited opportunities to monitor behaviour *before* as well as after TV's arrival.

Thus, despite claims to the contrary - particularly in the USA - links between TV viewing and viewers' behaviour remain unproven. Additionally, the equivocal nature of the TV-viewing effects' literature suggests that a simple, causal model for understanding TV effects is inappropriate and inadequate. A plenitude of personal and contextual characteristics most likely interacts with TV exposure to determine the nature and strength of any viewing effects (Wilson et al., 1998). Laboratory studies offer empirical support to this reasoning. Bandura, Ross and Ross (1961), for instance, demonstrated that children can learn anti-social behaviours from their viewing without performing them. Whether or not this learning was performed was determined by environmental factors. In a similar vein, Potts, Huston and Wright (1986) found that particular environmental factors, or cues, could expunge the influences of exposure to TV violence (although, in terms of the *performance* of, not the *learning* of, violent behaviours). Sawin (1990), too, highlighted the importance of situational variables as modifiers of any TV viewing effects. Hence, differential research outcomes may well reflect variations in context.

Regrettably, few possibilities have arisen to test this model in a naturalistic setting that allows the monitoring of behaviour both *prior* to as well as following the availability of TV. An ideal setting within which to help redress this paucity is a community with little or no exposure to broadcast television (yet soon to experience TV for the first time), which is characterized by low rates of anti-social behaviour and a social milieu which is vigilant over youngsters' behaviour. If the milieu remains influential, then TV's availability should have minimal impact upon social behaviour. Fortuitously, the island of St Helena offered such a setting and thus a rare opportunity to test this model.

St Helena, a UK colony in the South Atlantic Ocean, is among the world's most isolated populated islands. It is without daily newspapers, a cinema, an airport (although debates are now on-going to determine if one is to become operational in 2003) and, until recently, broadcast TV. TV (CNN) was beamed to St Helena in March 1995. Since then the network has expanded to incorporate 3 channels with the following services: M-Net Brochure (KTV, Movie Magic, BBC World Service), Discovery and Supersport.

The introduction of broadcast television was a major event on the island, whose only televisual experience had been through video. The video arrived in 1979. The island's 1987 census showed 29% of households owned one or more video sets. Whilst financial circumstances can limit the widespread availability of the sets, so can the non-availability of electricity. At that time, many households, particularly those in more remote and less accessible regions of the island, were unconnected to a main electricity supply. Even so, in 1994 (the year before broadcast television's availability) around 30% of 3- to 4-year-old children were watching videos (mostly cartoons) for about 11 minutes a day. The new televisual experience provided an unusual opportunity for a naturalistic investigation to monitor children's social behaviour across the advent of broadcast television.

The island's education system is well developed. During the seven-year span of the study, a number of changes have taken place. In 1993 there were six first schools (for 4- to 8-year-olds), three middle schools (8-12) and a secondary school (12+). Nursery provision was attached to each first school, and became available to all children when they were three and a half years old (in exceptional circumstances, younger children could be admitted). Although not obligatory, almost all children attended. By 2000, there were three first schools (for 4- to 7-year-olds), two middle schools (7-11), one amalgamated first/ middle school and one comprehensive school (12+). Nursery class provision became available at the beginning of the academic year for children whose fourth birthday occurred in that year.

METHOD

Participants

Participants in this study were drawn from nursery cohorts in St Helena in 1993 and 2000. As the admissions policies for the two cohorts were different - and, consequently, the cohorts' age-ranges were disparate - only similarly aged children were drawn from the 1993 and 2000 cohorts. Children were aged between 38 and 52 months in 1993 (23 boys, 27 girls) and 2000 (18 boys, 17 girls). Mean ages for the samples were 3 years 5 months (1993); and 3 years 7 months (2000).

Instrument

The Pre-school Behaviour Checklist (PBCL) was used by nursery teachers to rate their children's behaviour in 1993 and 2000. The PBCL (McGuire & Richman, 1988) includes 22 items on emotional functioning, conduct problems, social relations and concentration (see TABLE 1). Two items refer to speech and language functioning. Most items allow rating choices between three behavioural descriptions, scoring 0 (no problem), 1 (possible problem) or 2 (definite problem). The maximum score obtainable is 44.

McGuire and Richman (1988) provide reliability and validity data. Test-retest reliability is given as .88 ($p < .0001$) and internal consistency was measured using Cronbach's alpha (.83). **Validity was established at 83% through comparing PBCL scores with ratings from individual observers and interviews with staff. Do you mean there was an 83% agreement rate between PBCL scores and ratings from individual observers etc? Please clarify.**

TABLE 1 PBCL Items

	Item Description		Item Description
1.	Activity level	12.	Whines
2.	Not liked by others (peers)	13.	Is sensitive
3.	Wets	14.	Fights
4.	Soils	15.	Wanders aimlessly
5.	Has poor concentration	16.	Interferes with others
6.	Is difficult to manage	17.	Is miserable
7.	Demands attention	18.	Teases others
8.	Speaks unclearly	19.	Is withdrawn from staff
9.	Is reluctant to speak to others	20.	Is destructive
10.	Has temper tantrums	21.	Is fearful
11.	Is not sociable with peers	22.	Has habits (e.g. sucking & biting)

Procedure

In November 1993 (16 months prior to TV's availability) and November 2000 (nearly six years after TV's arrival) nursery class teachers completed the PBCL for children in their classes. Six teachers undertook ratings in 1993 and four teachers in 2000. All raters were female. A researcher marked completed checklists. In the absence of a directional hypothesis, data from the two cohorts with genders separated were subjected to two-tailed t tests. The t scores were calculated on the assumption of unequal variances.

For the 2000 cohort a random sample of pupils (n=14) was rated again by nursery class assistants (N = 4) to provide an inter-rater reliability check. This was calculated by dividing the number of rating agreements by the total number of ratings (Hall, 1974).

RESULTS

Individual PBCL total scores for boys and girls in both cohorts were positively skewed. PBCL total mean scores were calculated for boys (4.6) and girls (3.4) for the 1993 sample. Cronbach's α coefficient was .85. For the 2000 sample, total mean scores were 5.7 for boys and 3.9 for girls. Cronbach's α coefficient was .83. No significant differences on total mean scores were found between the two rating occasions for either boys, $t_{(30)} = 0.76$, $p > 0.05$, or girls $t_{(33)} = 0.42$, $p > 0.05$.

TABLE 2 Item Means for Boys in 1993 and 2000

	Boys				
	1993		2000		
ITEM	*M*	*SD*	*M*	*SD*	*p*
1. Activity level	0.48	0.51	0.28	0.58	.25
2. Not liked by peers	0.09	0.42	0.00	0.00	.33
3. Wets	0.04	0.21	0.11	0.47	.58
4. Soils	0.00	0.00	0.11	0.47	.33
5. Has poor concentration	0.52	0.51	0.94	0.73	.05*
6. Is difficult to manage	0.39	0.58	0.22	0.43	.29
7. Demands attention	0.35	0.49	0.50	0.62	.40
8. Speaks unclearly	0.30	0.47	0.50	0.62	.27
9. Reluctant to speak	0.04	0.21	0.22	0.43	.12
10. Has temper tantrums	0.04	0.21	0.17	0.38	.23
11. Not sociable with peers	0.13	0.34	0.28	0.46	.27
12. Whines	0.26	0.45	0.50	0.62	.18
13. Sensitive	0.17	0.49	0.28	0.46	.49
14. Fights	0.44	0.51	0.28	0.58	.37
15. Aimless wandering	0.04	0.21	0.39	0.50	.01**
16. Interferes with others	0.48	0.51	0.50	0.51	.89
17. Miserable	0.00	0.00	0.06	0.24	.33
18. Teasing others	0.39	0.50	0.11	0.32	.04*
19. Withdrawn from staff	0.09	0.29	0.06	0.24	.70
20. Destructive	0.30	0.47	0.06	0.24	.03*
21. Fearful	0.04	0.21	0.00	0.00	.33
22. Habits	0.00	0.00	0.11	0.32	.16
Total score on assessment	4.61	3.54	5.67	4.97	.45

T scores have been calculated on the assumption of unequal variances.

* significant at the .05 level

** significant at the .01 level

Comparisons between the PBCL item mean scores on the two rating occasions, showed few significant changes on the 22 comparisons for both genders (see Table 2). On the two occasions, teachers tended to rate children's behaviour in similar ways. With boys' and girls' scores separated, comparisons between the two sets of mean scores showed only six significant shifts. Boys attending nursery provision in 2000 were rated as having lower concentration levels $t\ (_{29}) = 2.10$, $p<0.05$, and as being more likely to 'wander aimlessly' $t\ (_{22}) = 2.74$, $p<0.01$. However, they were less likely to be either 'destructive' $t\ (_{34}) = 2.21$, $p<0.05$, or to 'taunt and tease others' $t\ (_{38}) = 2.17$, $p<0.05$. Girls in the 2000 sample were less likely to 'whine' $t(_{40}) = 2.19$, $p<0.05$, and were more 'reluctant to speak to others' $t(_{18}) = 2.20$, $p< 0.05$.

Inter-rater reliability for the 2000 nursery teachers' and nursery assistants' ratings was .83.

TABLE 3 Item Means for Girls in 1993 and 2000

	1993		2000		
ITEM	*M*	*SD*	*M*	*SD*	*p*
1. Activity level	0.19	0.40	0.44	0.73	.22
2. Not liked by peers	0.04	0.19	0.13	0.34	.36
3. Wets	0.00	0.00	0.00	0.00	/
4. Soils	0.00	0.00	0.00	0.00	/
5. Has poor concentration	0.37	0.57	0.69	0.79	.17
6. Is difficult to manage	0.15	0.36	0.06	0.25	.37
7. Demands attention	0.41	0.50	0.19	0.40	.12
8. Speaks unclearly	0.19	0.40	0.31	0.48	.38
9. Reluctant to speak	0.04	0.19	0.31	0.48	.04*
10. Has temper tantrums	0.19	0.40	0.13	0.34	.60
11. Not sociable with peers	0.19	0.40	0.25	0.45	.64
12. Whines	0.41	0.50	0.13	0.34	.03*
13. Sensitive	0.19	0.48	0.13	0.34	.64
14. Fights	0.04	0.19	0.06	0.25	.73
15. Aimless wandering	0.11	0.32	0.31	0.48	.15
16. Interferes with others	0.26	0.53	0.13	0.34	.32
17. Miserable	0.07	0.27	0.00	0.00	.16
18. Teasing others	0.30	0.61	0.13	0.34	.24
19. Withdrawn from staff	0.07	0.27	0.13	0.34	.61
20. Destructive	0.11	0.32	0.06	0.25	.58
21. Fearful	0.04	0.19	0.19	0.54	.30
22. Habits	0.04	0.19	0.13	0.34	.36
Total score on assessment	3.37	3.87	3.88	3.72	.68

T scores have been calculated on the assumption of unequal variances.

* significant at the 0.05 level

** significant at the 0.01 level

DISCUSSION

Perhaps the most important point in terms of the study's contribution to the global debate on the effects of TV watching on children's behaviour is this: There is no evidence whatsoever that the behaviour of young children in St Helena after nearly six years of exposure to TV has become more anti-social or more aggressive. We return below to a consideration of why our results run counter to so many others' findings, but in a word the key seems to be – context.

Of the six differences that the current study did discover between the 1993 and 2000 nursery cohorts, one stands out as the most robust, namely, that boys' concentration levels are rated as being significantly lower than in 1993. Its robustness derives from the fact that the other two comparisons carried out between the 1993 and subsequent nursery cohorts have yielded a similar picture (e.g. Charlton, Abrahams, & Jones, 1995; Charlton, Panting, Coles, & Hannan., 1999). The explanation may be, as Singer and Singer (1986) have argued, that high-action and fast-paced TV programmes can induce a dependence upon external stimuli that can lead to "increased restlessness in those children who have not already acquired play skills" (p.109). Thus, frequent exposure to fast-moving and action-packed TV (e.g. cartoons) may have adversely affected the 2000 cohort boys' concentration. Additionally, it has been found that children who are heavy viewers of cartoons are rated as low in enthusiasm by classroom teachers (Singer , 1982). However, in this study the finding that girls' concentration did not fall appears to weaken this argument. Further empirical work is planned to examine this apparent anomaly.

Even so, the explanation for the gender difference in concentration between the cohorts may also be linked to a wider, albeit related, phenomenon. Specifically, over this period (the end of the 1990s) boys everywhere have been increasingly exposed to video and electronic games (arguably more than girls). Many of these games have the characteristics which Singer and Singer highlight (high action, fast-paced, etc.). Hence, the explanation for the boys' lowered concentration levels may be related to TV-watching in so far as this shares with video games and other 'action-packed' modern technology the characteristics suggested by Singer and Singer. There is a further, arguably more plausible (but not mutually exclusive), possibility. Maybe a generation of young children (especially boys) with high exposure to the kind of instant gratification given by electronic games of the 'action' variety will find it increasingly difficult to apply their concentration to the more reflective activities typical of the classroom.

Gender differences of the above kind might, of course, have been affected by the gender of the raters, all of whom were female. However, in most school systems around the world, the teachers of the youngest children tend to be female and this was the case in St Helena. It is therefore possible that their ratings may have influenced the nature or extent of any sex differences in the results. We have no way of assessing this. However, this situation was a constant and therefore sheds no light on the changes in results over time. Future research will consider this factor as far as possible and also the possibility (remote, perhaps) that the exposure to broadcast television of the raters could have changed their perception of the behaviour of their pupils over time.

The other five differences between the 1993 and 2000 nursery cohorts are less robust, because they do not mirror differences found in the earlier cohort studies. For example, in an earlier study of 1993/1998 comparisons, the boys not only showed poorer concentration but they were more likely to be rated by their teachers as 'whining' and 'fearful' (e.g. Charlton, Abrahams, & Jones, 1995; Charlton, Panting, Coles, & Hannan, 1999). In the present study, these differences are not in evidence (at least at the level of statistical significance). On the other hand, the 1993/1998 comparisons of the girls revealed no statistically significant differences at all, whereas the 2000 cohort of girls is rated as less whining.

On these other five fronts, therefore, a closer analysis of the characteristics of the 2000 cohort in particular is being undertaken – for later publication. For example, the Island's policy in relation to First Schools and Nursery Classes changed in 2000 and it may be that some of the consequences of that policy change (on intake, age distribution, etc.) could have affected the findings.

Nevertheless, the major findings from the present analyses complement well and reinforce other findings in the St Helena study. For instance, Charlton, Panting, Davie, Coles and Whitmarsh (2000) video-recorded and then coded children's playground behaviour before and each subsequent year following the availability of TV, and failed to detect any increases in anti-social behaviour across that time period. However, the children's behaviour under investigation in that inquiry used an aggregate measure of youngsters' playground behaviour that was unconnected to individuals' viewing experiences. This limitation was corrected in the Gunter, Charlton, Panting and Coles' (2000) study, where they searched for links between children's anti-social behaviour and the amount of viewing undertaken as well as the amount of violence viewed by individuals. They found that viewers' behaviour was not significantly correlated to either the amount of viewing undertaken or the amount of violence watched.

The findings from three separate and independent measures comparing children's social behaviour on pre- and post-TV occasions failed to uncover any evidence to support widespread claims that TV viewing *per se* encourages anti-social behaviour. Equally important, results from another study in the St Helena project have suggested a way of accounting for this apparent contradiction between findings from this study and those found elsewhere.

Thus, one explanation for the tendency of the St Helenian children's good behaviour to remain relatively unaffected across the availability of TV viewing derives from the notion that any TV-viewing effects (adverse or otherwise) can be expunged or moderated by situational or contextual factors. Evidence to support this explanation has emerged from a focus group discussion involving older secondary school students on the island. The conclusion was that it is difficult to indulge in anti-social behaviour on St Helena: "because everyone watches you ... everyone knows you... You've just got to behave." (Charlton, & O'Bey, 1997). The students' remarks highlighted family and community social controls - or 'guardianship' behaviours - which appeared to be more influential in shaping children's behaviour than simple exposure to TV. This suggests that particular environmental factors can expunge or moderate any influences of exposure to TV violence (although, in terms of the expression of, not the learning of, violent behaviours).

The students' comments suggested one plausible explanation for the continuance of good behaviour even after the introduction of TV. In their discussions, they referred to a 'neighbourhood watch' on St Helena; a kind of informal pastoral network. Their comments implied that this watchfulness encouraged youngsters to become answerable to others (i.e., to give reckonings or explanations for their behaviour). In tandem, this watchfulness and the answerability linked to it, appeared to underpin good social behaviour by establishing common expectations for behaviour across homes, schools, and community. In other words, children's behaviour was shaped largely by environmental (situational) influences or cues, which were both consistent and beneficial. This thinking is not new. It has become a central tenet within most social learning theories. For example, Mischel (1977) expounded the power of situational determinants, and Rotter's (1966) social learning theory highlighted the import of the 'psychological situation' in shaping behaviour. Moreover, such thinking was tested empirically by Potts, Huston and Wright (1986). They studied the influence upon young boys' behaviour of a systematic manipulation of situational cues (in this case, aggressively and pro-socially cued toys) after exposing the youngsters to varied TV stimuli (i.e., high and low action, high and low violence). Their results showed that particular environmental attributes can expunge the influences of exposure to TV violence (in terms of the expression of, not the learning of, violent behaviours). The nature of the above thinking and findings suggests the conclusion that in watchful and caring environments, dominant determinants of behaviour can exclude TV.

The influence of environmental factors has also been stressed elsewhere. Social controls have been noted at work in small - and mainly rural - communities. Valentine (1997), for example, maintained that in communities where most people are known to most others, feelings of being watched over are commonplace because of the 'eyes on the street' (p.144) that keep young and old alike under surveillance. Likewise, Cohen (1982) talks of individuality giving way to communality, and Jones (1999, p.9) refers to the 'power of gossip' in establishing conformity. Another (related) perspective argues for the importance of social capital (e.g. the extent to which children have others to draw upon for support and help) in determining their behaviour (Amato & Booth, 2000). Comment of this kind draws attention to social controls of one type or another with the potential to affect, favourably or otherwise, individuals' behaviour.

These kinds of explanations and models are perhaps strengthened in relation to the present study by reported crime rates on St. Helena, which have remained low across nearly six years of broadcast TV; indeed crime figures (including violent crime) were most recently reported as being halved ("Half Year Figures." Vol. 2: Summary report. 2000). Again, this is an area that merits more detailed examination.

The major finding, thus far, from this study and the wider St Helena TV project, is that in the short-term (i.e. after nearly six years of TV) - contrary to most other research findings - TV viewing does not unequivocally or inevitably influence children's social behaviour. Although a few significant (and interesting – not to say, puzzling) differences have emerged between children's pre- and post-TV behaviour, none of these was related to aggressive or anti-social tendencies.

REFERENCES

Amato, P. R., & Booth, A. (2000). *A Generation At Risk*. Harvard University Press, Cambridge, Mass.: Harvard University Press.

Bandura, A., Ross, D., & Ross, S.A. (1961). Transmission of aggression through imitation of aggressive models. *Journal of Abnormal and Social Psychology, 63*, 575-582.

Boyatzis, C., Matillo, G., & Nesbitt, K. M. (1995). Effects of 'The Mighty Morphin Power Rangers' on children's aggression with peers. *Child Study Journal, 25*, 45-55.

Charlton, T., Abrahams, M., & Jones, K. (1995). Prevalence rates of emotional and behavioral disorder among nursery class children in St. Helena, South Atlantic: An epidemiological study. *Journal of Social Behavior and Personality, 10*, 273-280.

Charlton, T., & O'Bey, S. (1997). Links between television and behaviour: Students' perceptions of television's impact in St Helena, South Atlantic. *Support for Learning, 12*, 3, 130-136.

Charlton, T., Panting, C., Coles, D., & Hannan, A. (1999). Behaviour of nursery class children before and after the availability of broadcast television: A naturalistic study of two cohorts in a remote community. *Journal of Social Behaviour and Personality, 14*, 3, 315-324.

Charlton, T., Panting, C., Davie, R., Coles, D., & Whitmarsh, L. (2000). Children's playground behaviour across five years of broadcast television: A naturalistic study in a remote community. *Emotional and Behavioural Difficulties, 5*, 4, 3-11.

Cohen, A.P. (1982). *Belonging: Identity and Social Organisation In British Cultures.* Manchester: Manchester University Press.

Coldevin, G., & Wilson, T. C. (1985). Effects of a decade of satellite television in the Canadian arctic; Euro-Canadian and Inuit adolescents compared. *Journal of Cross-Cultural Psychology, 16*, 3, 329-354.

Granzberg, G. (1985). Television and self-concept formation in developing areas: The central Algonkian experience. *Journal of Cross-Cultural Psychology, 16*, 3, 313-328.

Gunter, B., Charlton, T., Panting, C., & Coles, D. (2000). The impact of television on children's anti-social behavior in a television community. *Child Study Journal, 30*, 2, 65-90.

Hall, R.V. (1974). *The Measurement of Behaviour.* Lawrence, KS: H & H Enterprises.

Jones, G. (1999). "The same people in the same places?": Socio-spatial identities and migration in youth. *Sociology, 33*, 1-22.

McGuire, J., & Richman, N. (1988). *The Pre-School Behaviour Checklist.* NFER-Nelson: Windsor.

Mischel, W. (1970). Sex differences. In. P. Mussen (Ed.). *Carmichael's Manual of Child Psychology* (Vol. 2) New York: Wiley.

Potts, R., Huston, A. C., & Wright, J. C. (1986). The effects of television form and violent content on boys' attention and social behaviour. *Journal of Experimental Child Psychology, 41*, 1-17.

Rotter, J.B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs, 80*, (Whole No. 609).

Sawin, D. B. (1990). Aggressive behaviour among children in small playground settings with violent television. *Advances in Learning and Behavioural Difficulties, 6*, 157-177.

Singer, D. (1982). Television and the developing imagination of the child. In. Pearl, D., Bouthilet, L., & Lazar, J. (Eds.). *Television and Behaviour: Vol. 2: Summary Report. Ten years of scientific progress and implications for the eighties*, Washington, DC.: Government Printing Office.

Singer, J. L., & Singer, D. G. (1986). Family experiences and television viewing as predictors of children's imagination, restlessness and aggression. *Journal of Social Issues, 42*, 107-124.

Smith, S. L., Nathanson, A. I., & Wilson, B. (May, 1999). *Violence in Prime Time: An Analysis of the Amount and Context of Aggression.* Paper presented at the Annual Conference of the International Communications Association, San Francisco, California.

Wiegman, G., Kuttschreuter, M., & Baarda, B. (1992). A longitudinal study of the affects of television viewing on aggressive and pro-social behaviours. *British Journal of Social Psychology, 31*, 147-164.

Williams, T. M. (1986). *The Impact Of television. A Natural Experiment In Three Communities*. Academic Press Inc.

Wilson, B. J., Kunkel, D., Linz, D., Potter, J., Donnerstein, E., Smith, S.L., Blumenthal, E., Berry, M., & Federman, J. (1998). The nature and context of violence on American television. In U. Carlsson & C. Von Feilitzen (Eds.). *Children and Media Violence*. Goteborg, Sweden: UNESCO.

Valentine, G. (1997). A safe place to grow up? Parenting, perceptions of children's safety and the rural idyll. *Journal of Rural Studies, 13*, 137-148.

The Self-Reported Psychological Well-Being of Celebrity Worshippers

John Maltby
Sheffield Hallam University, England, U.K.

Lynn E. McCutcheon
DeVry Institute of Technology – Orlando

Diane D. Ashe
Valencia Community College

James Houran
Southern Illinois University School of Medicine

The absorption-addiction model of celebrity worship suggests that celebrity adoration is accompanied by a poor psychological well-being. To test the validity of this model we administered the Celebrity Attitude Scale (CAS) and the General Health Questionnaire (GHQ-28) to a convenience sample of 126 men and 181 women from the United Kingdom. Generally the correlations between the variables were consistent with predictions. Further tests provided convergent evidence for the absorption-addiction model. Specifically, the Entertainment-Social sub-scale of the CAS accounted for unique variance in Social Dysfunction, and Depressive Symptoms, while the Intense-Personal sub-scale accounted for unique variance in Depression and Anxiety scores. We argue that the positive relationship between Celebrity worship and poorer psychological well-being results from (failed) attempts to escape, cope, or enhance one's daily life.

Two competing characterizations of *celebrity worship* have dominated the social scientific literature. One of these suggests that celebrity worship is an expression of pathology. Anecdotal evidence is usually cited for this position, and includes fans' obsession with certain television programmes (Jindra, 1994), the attempts to harm celebrities (Caughey, 1978; Giles, 2000; Schickel, 1985), the threatening letters they mail to movie and television stars (Dietz, Matthews, Van Duyne, Martell, Parry, Stewart, Warren, & Crowder, 1991), and fans' apparent confusion between celebrities' fictional roles and their real lives (Caughey, 1978; Giles, 2000; Klapp, 1962).

Author info: Correspondence should be sent to: Dr. John Maltby, School of Health, Sheffield Hallam University, Sheffield, England, S10 2BP.

North American Journal of Psychology, 2001, Vol. 3, No. 3, 441-452.
© NAJP

Other commentators offer a different view of celebrity worship. For example, Jenkins (1992) and Jenson (1992) proposed a complicated cultural interaction between fans and celebrity worship. They argue that, rather than pointing to pathology, we need to emphasise the active, positive role that these fans undertake in creating and maintaining social networks around their favorite celebrities. As such, celebrity worship might reflect the development of a deeper appreciation and enthusiasm for particular people and their talents (Giles, 2000; Jenkins, 1992; Jenson, 1992).

Except for the initial reports of interview data (Giles, 2000), there has been little empirical evidence to support either characterization of celebrity worship. Furthermore, commentators make little anecdotal distinction in the literature between examples of extreme behaviour on the part of fans and how being involved in celebrity worship might have other (even positive) consequences for people. Indeed, at least three different hypotheses of how celebrity worship might impact an individual's psychological well being can be formulated. We review these next.

Speculations on the Relation of Celebrity Worship to Psychological Well-Being

Our introduction provides the basis for three viewpoints concerning the relationship between celebrity worship and psychological well being. First, celebrity worship can be seen as detrimental to the individual. Here, celebrity worship is typically characterised by a single individual obsessed by a particular celebrity (Dietz, et al. 1991; Giles, 2000), the "worshipper" is usually alone in the worshipping activities (and sometimes lonely in their activity; Rubin, Perse, & Powell, 1985), and the individual demonstrates some aspect of pathology. In other words, such an individual would have poor psychological functioning and well-being.

Recently, McCutcheon, Lange, and Houran (in press) provided strong empirical evidence for the general pathological view of celebrity worship. Using statistical tests for dimensionality (Rasch modeling and tests for differential item functioning), these authors analysed a large set of questions about celebrity worshipping behaviours. Several novel findings emerged. Specifically, items not discarded due to age or gender bias formed a unidimensional hierarchy that applied to many types of celebrities (e.g., actors, musicians, sports figures). Moreover, the analyses did not reveal any compelling evidence to conceptualise celebrity worship as having distinct non-pathological and pathological forms.

Instead, the "Rasch" nature of the items defines celebrity worship as consisting of three qualitatively different stages on a single continuum

(McCutcheon et al., in press). In particular, low worship, as operationalized by low scores on the CAS, involves individualistic behaviours such as watching and reading about a celebrity. At slightly higher levels, celebrity worship takes on a social character. Lastly, the highest levels are characterized by a mixture of empathy with the celebrity's successes and failures, over-identification with the celebrity, and compulsive behaviours, as well as obsession with details of the celebrity's life. Based on these findings, the authors (McCutcheon et al., in press) proposed a model of celebrity worship based on *psychological absorption*. This absorption is promoted by an attempt to establish a sense of identity and fulfilment (leading to delusions of actual relationships with celebrities) and later reinforced by *addiction* (fostering the need for progressively stronger involvement to feel connected with the celebrity). The factor analysis (Maltby, Houran, Lange, Ashe, & McCutcheon, in press) of a modified version of McCutcheon et al.'s (in press) original questionnaire likewise validated this basic notion of three expressions of celebrity worship – an entertainment/ social dimension, intense personal feelings toward a celebrity, and borderline pathology.

However, a second and contrasting view of celebrity worship argues that this behaviour is potentially beneficial to the individual, provided that the individual is participating in a social network of fans. In this respect, celebrity worship involves nothing more than an individual attending conferences or sharing information and experiences with friends (or with individuals on the Internet). Such behaviours might promote productive social relationships and serve as a psychological buffer against every-day stressors. According to this logic, the consequences of socially based adulation can be beneficial to the individual.

Finally, we must consider the possibility that there is no simple relationship between celebrity worship and a person's psychological well being. That is, there is a complex cultural interaction between the individual and celebrity worship (Jenkins, 1992; Jenson, 1992), based on an appreciation of ability and the medium, which is more dynamic than is captured in a correlational analysis. Therefore, this perspective might suggest that there is no theoretical reason to hypothesise a direct relationship between celebrity and psychological well-being.

Testing the Relation of Celebrity Worship to Psychological Well-Being

The theories outlined above characterise celebrity worshippers differently. While one viewpoint emphasises the individual as obsessive, alone, and showing poorer psychological well-being, another viewpoint emphasises celebrity worship as beneficial, with the individual integrated into a social world, gaining social skills and support from participating in celebrity worship. Thus, it would seem prudent to make such distinctions when examining psychological and behavioural correlates of celebrity worship (or its hypothesized stages).

Recent research with the *Celebrity Attitude Scale* (McCutcheon et al., in press) among UK samples (Maltby et al., in press) suggests that this scale can be used to reflect distinctions among the three different stages or aspects of celebrity worship.[1] In particular, the first factor is an "*Entertainment-social*" factor comprising items like "My friends and I like to discuss what my favourite celebrity has done [item 4]", and "Learning the life story of my favourite celebrity is a lot of fun [item 15]." This subscale reflects social aspects to Celebrity worship and is consistent with Stever's (1991) observation that fans were primarily attracted to their favourite celebrity because of their perceived ability to entertain. Rubin and McHugh (1987) made a similar observation. Both of these studies referred to this as "task attraction."

The second factor is labelled "*Intense-personal*," and is exemplified by items like "I consider my favourite celebrity to be my soul mate [item 10]", and "I have frequent thoughts about my celebrity, even when I don't want to [item 11]". It can be suggested that this scale measures individuals' intensive and compulsive feelings around the celebrity, reflective of the obsessional tendencies of fans often referred to in the literature (Dietz, et al., 1991).

The third factor, labeled "*Borderline pathological,*" is exemplified by items like "If someone gave me several thousand dollars to do with as I please, I would consider spending it on a personal possession (like a napkin or paper plate) once used by my favourite celebrity [item 20]," and "If I were lucky enough to meet my favourite celebrity, and he/she asked me to do something illegal as a favour I would probably do it [item 17]." The sub-scale seems to reflect some of the extreme attitudes individuals may hold regarding their favourite celebrity.

The *General Health Questionnaire - 28* is an easily administered, self report, screening device designed to reveal symptoms of poor psychological health. Factor one was labelled Somatic Symptoms; factor two is anxiety and insomnia, factor three is social dysfunction and four is severe depression. Both the 28-item version and the 60-item version from which it is derived have been shown to possess good psychometric properties (Goldberg & Williams, 1991).

---

[1] It remains to be seen whether these three factors show a similar Rasch hierarchy (or progression) as reported in McCutcheon et al. (in press), but we feel it is encouraging that the Maltby et al. (in press) findings also suggested that celebrity worship involves three aspects that appear to coincide in content with the three hierarchial stages found by McCutcheon et al. (in press).

These two measures allow us to test the three theories reviewed. Specifically, the aim of the present study was to examine the relationship between the three aspects of celebrity worship (Entertainment-Social, Intense-personal and Border- line-Pathological) and psychological well being. Specifically, we wished to determine whether different psychological variables motivate the three different expressions of celebrity worship as predicted by McCutcheon et al.'s (in press) model of celebrity worship.

METHOD

Measures

1. *The Celebrity Attitude Scale* (CAS). The CAS is a 34-item Likert scale with "strongly agree" equal to 5 and "strongly disagree" equal to 1. From analysis reported in Maltby, et al. (in press), three subscales were formed from 23 of the items; Entertainment/Social; Intense-Personal, and Borderline pathological.

2. *The General Health Questionnaire – 28* (Goldberg & Williams, 1991). Each of these scales comprises 7-item measures of depressive symptoms, anxiety symptoms, social dysfunction and somatic symptoms. Responses are scored on a 4-point scale ranging from (0) Better than usual, (1) Same as Usual, (2) Worse than Usual, (3) Much Worse than usual. Higher scores indicate a greater degree of self-reporting of each of the symptoms.

Participants

A convenience sample of 126 men (M_{age} = 26.97 years, SD = 6.8, range = 18 to 48) and 181 women (M_{age} = 27.67 years, SD = 7.9, range = 18 to 48) was recruited from a number of workplaces and community groups in the South Yorkshire region of England. The most frequently cited demographic categories were Caucasian (n = 220), single (n = 144), employed (n = 196), and an "O" educational level or its equivalent (n = 85), which approximates a high school education in the United States. The number of recruits who declined to participate or failed to complete the scales was 22. Data using the Celebrity Attitude Scale here have been reported elsewhere (Maltby et al. in press), however, the correlations reported among all the variables measured in this study have not been reported elsewhere.

Procedure

Participants were told their responses were confidential and that the purpose of the study was to examine a number of psychological factors that may be related to individuals' interest in famous people. The scales were presented to participants in different orders to reduce the probability of a systematic order effect. Participants completed the scales in a single session in small groups at their workplace or community meetings.

RESULTS

Preliminaries.

Table 1 shows mean scores and standard deviations for each measure. There were no significant sex differences for any of the scales, and none of the *t* values were higher than .54. Therefore, in the following analyses, data for men and women were combined.

TABLE 1: Means (standard deviations) for all the scales by sex.

Scales	Men (n=126)	Women (n=181)
1. Entertainment-Social	15.81 (05.7)	16.07 (06.0)
2. Intense-Personal	20.31 (07.3)	20.53 (06.8)
3. Borderline Pathological	07.65 (02.5)	07.63 (02.5)
4. Somatic Symptoms	03.76 (03.2)	03.66 (03.0)
5. Anxiety	01.90 (02.5)	01.90 (02.0)
6. Social Dysfunction	03.22 (03.2)	03.17 (02.9)
7. Depression	01.85 (01.8)	02.00 (02.2)

Main Findings

Table 2 gives the Pearson product moment correlations between all the variables. All of the Celebrity Attitude Subscales significantly and positively correlated with one another. Scores for the Entertainment-Social Celebrity Attitude subscale correlated positively and significantly with Anxiety, Social Dysfunction, and Depressive Symptoms, and both the Intense-Personal and Borderline Pathological Celebrity Attitude subscales correlated significantly and positively with Anxiety and Depressive Symptoms.

Table 3 depicts these relationships in the context of the hypothesized progression of the stages of celebrity worship (cf. McCutcheon et al., in press). Specifically, we propose that celebrity worship begins at the Entertainment/Social stage, progresses to an Intense-Personal dimension, and finally reaches the Borderline pathological domain. Assuming this progression is valid, we find that the initial stage of celebrity worship is marked by a component of social dysfunction, as would be expected if this stage is motivated by psychological absorption, i.e., a yearning to soothe "the empty self" due to the worshipper's likely introverted nature and lack of meaningful relationships (Willis, 1972; Szymanski, 1977; Stever, 1995; Meloy, 1998). The next stage (Intense-Personal domain) involves anxiety and depression as well, but now takes on a somatic quality that we hypothesize reflects the now addictive component to celebrity worship. Finally, the Borderline-Pathology domain shows an exclusive mix of anxiety and depressive symptomatology, which parallels the progressive absorption-addiction model of celebrity worship that we proposed earlier (McCutcheon et al., in press).

TABLE 2: Pearson product-moment correlation coefficients between all the scales.

	1	2	3	4	5	6	7
1. Entertainment-Social	-						
2. Intense-Personal	.29**	-					
3. Borderline Pathological	43**	.37**	-				
4. Somatic Symptoms	-.01	.14*	.07	-			
5. Anxiety	.17**	.25**	.20**	.01	-		
6. Social Dysfunction	.22**	.02	-.01	-.03	-.03	-	
7. Depressive Symptoms	.26**	.22**	.21**	-.06	.28**	.18**	-

*$p<.05$, ** $p<.01$.

TABLE 3 Comparison of Pearson Correlations Between Indices of Psychological Well Being and the Three Hypothesized Stages of Celebrity Worship

	Stage 1: Entertainment-Social Domain	Stage 2: Intense-Personal Domain	Stage 3: Borderline-Pathological Domain
Somatic Symptoms	-.01	**.14**	.07
Anxiety	**.17**	**.25**	**.20**
Social Dysfunction	**.22**	.02	-.01
Depressive Symptoms	**.26**	**.22**	**.21**

Note: Highlighted coefficients visually indicate which psychological variables are most associated with each stage of celebrity worship.

To test the validity of these conclusions we conducted a standard multiple regression to examine which of the Celebrity Attitude sub-scales scores accounted for unique variance in Anxiety and Depression. We expected that depressive symptoms (including social dysfunction) are most indicative of the initial stages of celebrity worship (the absorption component due to a lack of identity) whereas anxiety would be manifested in the later stages (due to increasing thresholds to satisfy the addiction to the celebrity). Table 4 shows the results of two standard multiple regressions performed with each of the Celebrity Attitude subscales (Entertainment-Social; Intense-Personal, and Borderline-Pathological) used as independent variables, and the Anxiety and Depression used as dependent variables. Included in these tables are the unstandardised regression coefficients (B), the standardised regression coefficients (*B*), the semipartial correlations (sr^2), r, r^2 and adjusted r^2.

TABLE 4 Multiple Regression Analyses Using all Three Celebrity Attitude Sub-scales as Independent Predictors

Scales	B	Beta	Sr^2		B	Beta	Sr^2
Anxiety				Depression			
Entertain.-Social	.02	.08		Entertain.-Social	.07	.19	.04**
Intense-Personal	.06	.19	.04**	Intense-Personal	.04	.13	.02**
Borderline Path.	.08	.10		Borderline Path.	.07	.08	
		r^2	=.08			r^2	=.10
		Adj r^2	=.07			Adj r^2	=.09
		R	=.28			r	=.31

* p<.05; ** p<.01

For Anxiety, the regression statistic (r) was significantly different from zero ($F_{(3,303)} = 8.837$, $p<.001$, with the Intense-Personal subscale accounting for unique variance in Anxiety scores. For depressive symptoms, the regression statistic (r) was significantly different from zero ($F_{(3,303)} = 10.639$, $p<.001$), with the Entertainment-Social and Intense-Personal Celebrity Attitude subscale accounting for unique variance in depressive symptom scores.

DISCUSSION

The pattern of results is consistent with predictions from the absorption-addiction model of celebrity worship (McCutcheon et al., in press), and thus suggests that celebrity worship reflects poorer psychological well-being. This conclusion holds even for the initial stages of celebrity worship that do not appear pathological. For instance, the Entertainment-Social stage accounts for unique variance in Social Dysfunction and Depressive Symptoms, and celebrity worship for Intense-Personal reasons accounts for unique variance in Depression and Anxiety scores.

Using these relationships it is possible to refine our absorption-addiction model of celebrity worship (McCutcheon et al., in press). Superficially it may seem contradictory that celebrity worship for Entertainment-Social reasons is accompanied by Social Dysfunction.

However, if these significant positive correlations are placed in context of some of the items ("It is enjoyable just to be with others who like my favourite celebrity" [item 23] and, "One of the main reasons I maintain an interest in my favourite celebrity is that doing so gives me a temporary escape from life's problems" [item 8], then we suggest that this celebrity worship domain represents a person's attempt to establish an identity. Alternatively, it may be the only way the person is able to interact with others – a socially dysfunctional individual whose social life is solely based on worshipping a celebrity. This view is supported by the fact that the Entertainment-Social domain accounted for unique variance in the measure of depressive symptoms. As such, we would characterise high scorers on this domain as persons with little interest in anything apart from their chosen celebrity. Future research may find that individuals in this stage of celebrity worship use their interest in celebrities as a coping mechanism to deal with the stressors of daily life.

Scores on the Celebrity Attitude Intense-Personal domain account for unique variance in Anxiety, and to a lesser extent Depressive symptoms. The relationship between this aspect of Celebrity worship and Anxiety is perhaps less surprising given that some of the items of the scale "(I am obsessed by details of my favourite celebrity's life [item 3]") may reflect obsessional (i.e., addictive) tendencies, and anxiety is thought to be related to obsessionality (Sandler & Hazari, 1960). Therefore, this finding perhaps locates this type of celebrity worship securely within an Anxiety/Obsessionality domain. In addition, higher scores on the Intense-Personal domain were associated with higher scores on depressive symptoms. One explanation for this finding is that both anxiety and obsessional-neurotic traits are related to depression (Gaynes, Magruder, Burns and Broadhead, 1997; Kostanskil & Gullone, 1998; Maltby, Lewis & Day, 1999; Maltby, Lewis, & Hill, 1998; Wolfradt & Straube, 1998). Therefore, previous literature provides a context to understand why individuals experiencing Intense-Personal feelings in relation to their favourite celebrity also report depression. Further research might use measures of obsessionality, neuroticism, and other related constructs to further explore the Intense-personal aspect of Celebrity Worship.

Finally, scores on the Borderline-Pathological domain do not account for unique variance in any of the psychological well-being measures. Instead, any significant relationship found between this aspect of Celebrity Worship and the psychological variables might be better considered within the other two aspects of Celebrity Worship. Methodological limitations in our study also may have contributed to the fact that the Borderline-Pathological domain did not account for unique variance in poorer psychological well being. For example, the mental health measures we used do not cover those domains that might be expected with this type of pathological behaviour. The General Health

Questionnaire is designed for use among non-clinical populations, and therefore further research is needed to examine what particular aspects of psychological well being are most relevant to the Borderline-pathological domain. However, the present study provides no evidence that the Borderline-Pathological domain of celebrity worship is related to a richer psychological well being.

Further research is needed to explore the possible psychological processes that mediate our proposed absorption-addiction model of celebrity worship, which places this phenomenon in the context of (failed) attempts to escape, cope, or enrich one's everyday life. McCutcheon et al.'s (in press) model is a first approximation of a cohesive interpretation of the empirical research on celebrity worship. Future studies may suggest more viable models. Our understanding of the precise cognitive and emotional dynamics that underlie celebrity worship is in its infancy, but the available evidence indicates that celebrity worship is a behavioural expression of poor global psychological well being.

REFERENCES

Caughey, J. L. (1978). Artificial social relations in America. *American Quarterly*, Spring, 70-89.

Dietz, P. E., Matthews, D. B., Van Duyne, C., Martell, D. A., Parry, C. D., Stewart, T., Warren, J., & Crowder, J. D. (1991). Threatening and otherwise inappropriate letters to Hollywood celebrities. *Journal of Forensic Sciences, 36*, 185-209.

Gaynes, B.N., Magruder, K.M., Burns, B.J., & Broadhead, W.E. (1997). Persistaence of depressive illness in primary care patients with major depression: Is a coexisting anxiety disorder a risk factor? *International Journal of Psychiatry in Medicine, 27*, 307-308.

Giles, D. (2000). *Illusions of Immortality: A psychology of fame and celebrity.* London: MacMillan.

Goldberg, D., & Williams, P. (1991). *A user's guide to the general health questionnaire.* London: NFER Nelson.

Jenkins, H. (1992). *Textual Poachers: Television Fans and Participatory Culture.* New York: Routledge.

Jenson, J. (1992). Fandom as pathology: The consequences of charac- terization. In L. Lewis (Ed.) *The Adoring Audience: Fan Culture and Popular Media.* London: Routledge,

Jindra, M. (1994). Star-Trek fandom as a religious phenomenon. *Sociology of Religion, 55*, 27-51.

Klapp, O. E. (1962). *Heroes, Villains, and Fools.* Englewood Cliffs, NJ: Prentice-Hall.

Kostanskil, M. & Gullone, E. (1998). Adolescent body image dissatisfaction. Relationship with self-esteem, anxiety and depression controlling for body mass. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 39*, 255-262.

Maltby, J, Houran, J., Lange, R,. Ashe, D. & McCutcheon, L.E. (in press). Thou Shalt Worship No Other Gods - Unless They Are Celebrities: The relationship between celebrity worship and religious orientation. *Personality and Individual Differences.*

Maltby, J., Lewis, C.A. & Day, L. (1999). Religious orientation and psychological well-being: The role of the frequency of personal prayer. *British Journal of Health Psychology, 4*, 363-378.

Maltby, J., Lewis, C.A., & Hill, A.P. (1998). Oral Pessimism and depressive symptoms: A comparison with other correlates of depression. *British Journal of Medical Psychology, 71*, 195-200.

McCutcheon, L. E., Lange, R., & Houran, J. (in press). Conceptualization and measurement of celebrity worship. *British Journal of Psychology.*

Meloy, J.R. (Ed.) (1998). *The psychology of stalking: Clinical and forensic perspectives.* San Diego, CA: Academic Press.

Rubin, A.M., Perse, E., & Powell, R.A. (1985). Loneliness, parasocial interaction, and local television news viewing. *Human Communication Research, 12*, 155-180.

Rubin, R. B., & McHugh, M. P. (1987). Development of parasocial interaction relationships. *Journal of Broadcasting & Electronic Media, 31*, 279-292.

Sandler, J., & Hazari, A. (1960). The obsessional: On the psychological classification of obsessional character traits and symptoms. *British Journal of Medical Psychology, 33*, 113-122.

Schickel, R. (1985). *Intimate strangers. The culture of personality.* Garden City, NY: Doubleday.

Stever, G.S. (1991). The Celebrity Appeal Questionnaire. *Psychological Reports, 68*, 859-866.

Stever, G.S. (1995). Gender by type interaction effects in mass media subcultures. *Journal of Psychological Type, 32*, 3-22.

Szymanski, G. G. (1977). Celebrities and heroes as models of self-perception. *Journal of the Association for the Study of Perception, 12*, 8-11.

Willis, S. E. (1972). Falling in love with celebrities. *Sexual Behavior, 2*, 2-8.

Woldradt, U. & Straube, E.R. (1998). Factor structure of schizotypal traits among adolescents. *Personality and Individual Differences, 24*, 201-206.

BEST COPY AVAILABLE

Effects of a Health and Relationship Education Program on Drug Behaviors

Joseph Donnelly
Heather Ferraro
Carolyn Eadie
Montclair State University

Project C.A.R.E. is a community trial of the *Sex Can Wait* upper elementary and middle school curricula (Core-Gebhart, Hart, & Young, 1997), implemented in grades six, seven and eight at three urban middle schools in comparison to three control schools, all in lower socioeconomic status communities of northeastern New Jersey. The curriculum presents postponement of intercourse until after marriage as a necessary prerequisite for long-term satisfaction in a healthy and fulfilling life. In addition to factual information regarding sexual behavior, pregnancy, child rearing and their consequences, the curriculum emphasizes skills for goal setting, decision making, effective communication, and self-esteem enhancement areas commonly addressed by drug education curricula. Impact of exposure to the curriculum on use of alcohol, cigarettes and illicit drugs, therefore, seems to be a reasonable subject of investigation. Post-test comparisons revealed no differences between the groups in terms of drug attitudes or the use of alcohol, marijuana, or cigarettes. Students at comparison schools reported greater increases in use of stimulants and of cocaine/crack than did those attending intervention schools.

The two topics most often addressed in health education curricula are drugs and sex. Behavioral decisions made in these areas can be powerful determinants of overall health and well being. Both drug use and sexual activity necessitate decision-making, as both involve a degree of risk. Indeed, sexual activity among adolescents results in an estimated one million teen pregnancies per year in the U.S., 78% of which are unplanned (Guttmacher Institute, 1999). In addition, three million teens acquire a sexually transmitted disease each year (Guttmacher Institute, 1999). *Monitoring the Future*, a continuing study by the National Institute on Drug Abuse (NIDA of drug use trends among America's youth has found that 80% of all high school seniors have used alcohol; 65%have smoked cigarettes; 50% have used marijuana and 10% have

---

Author info: Correspondence should be sent to: Dr. Joseph Donnelly, Dept. of Health Professions, Montclair State U., Upper Montclair, NJ 07043.

North American Journal of Psychology, 2001, Vol. 3, No. 3, 453-462.
© NAJP

used cocaine (Johnston, O'Malley & Bachman, 2000). Besides having a profound impact on young people's health, substance use by young people has serious and costly consequences for society as well. The total cost of alcohol use by youth - including traffic crashes, violent crime, burns, drowning, suicide attempts, Fetal Alcohol Syndrome, alcohol poisoning and treatment - has been estimated to be more than $58 billion per year (Levy & Stewart, 1999).

The skills needed to successfully navigate choices related to both sex and drug use are the same - communication, self-esteem, resistance and refusal, and decision-making. Mayer (1999) maintains that the two major components of a successful substance abuse prevention program are "normative education" and resistance skills. Normative education involves teaching adolescents that using drugs is not acceptable and that everyone is not doing it. Resistance skills are about learning how to say no to drugs in different social situations.

A report called "Making the Grade: A Guide to School Drug Prevention Programs," identifies characteristics of a good school-based prevention program. It is one which promotes the development of refusal skills; helps students recognize internal and external pressures; teaches about the short- and long-term consequences of drug use; involves family and community; and uses culturally relevant materials (Drug Strategies, 1996).

Low self-esteem in children may make them vulnerable to drug abuse and other risk-taking behaviors as adolescents (Reasoner, 2000). In a study of 1,228 teens aged 14-17, Hofler & Lieb (1999) explored risk factors for progressive use of marijuana. Significant predictors of whether marijuana use progressed from "no use" to "regular use" were: family history of substance abuse, self-esteem, unconditional commitment to not using drugs, immediate availability of drugs, peer group use and previous history of nicotine and alcohol abuse. Low self-esteem and a lack of commitment to not use drugs were found to be powerful risk factors for use and "regular use" of marijuana.

Kawabata, Cross, Nishioka, and Shimai (1999) examined the relationship between self-esteem and smoking behavior among 2,090 Japanese students in grades 4-9. They found that students who had ever smoked showed significantly lower self-esteem in perceived cognitive and family relationship subscales than students who had never smoked. Based on their findings, the authors suggest that future smoking prevention programs targeted for early adolescence should be integrated into more comprehensive health education programs, which include self-esteem enhancement training.

Communication is another skill essential to negotiating choices about sex and drug use. A study reviewing peer-led drug prevention programs across North America found that the peer-led programs were superior to teacher-led programs in their effect on student use of alcohol, marijuana, and tobacco (Black & Tobler, 1998). The peer-led programs produced significant changes in cognition and attitude among middle school children. The same study found that interactive programs, involving face-to-face communication among peers, were most effective. In this approach, students learn from each other through role plays, modeling skills, and giving and receiving feedback. Peer mentors and mediators are also thought to benefit by gaining leadership and problem-solving skills, better communication, and an improved self-esteem (Humphries, 1999).

Many sexuality education curricula also focus on developing self-efficacy, self-esteem, resistance and refusal skills and social support (Donnelly, Goldfarb, Duncan, Young, Eadie & Castiglia,1999; Ferguson, 1998). Ferguson (1998) evaluated the effects of peer counseling in a culturally specific adolescent pregnancy program. Sixty-three African American females between the ages of 12 and 16 participated in the eight-week long program, which involved discussions and role-playing regarding peer pressure, decision-making, communication, and leadership skills, from an Afrocentric perspective. It was found that none of the participants who received peer counseling became pregnant after three months, although this outcome may be due to the fact that many of them had been sexually inactive before the intervention. Although it is difficult to assess the true effectiveness of the program for pregnancy prevention, participants did show an increase in their knowledge related to sex, reproduction, STDs and contraceptives.

Whitbeck, Yoder, Hoyt, and Conger (1999) used an event history analysis to examine predictors of early sexual intercourse for a sample of 457 adolescents in grades 8-10. They found that the primary predictors of early intercourse were age, opportunity, sexually permissive attitudes, association with delinquent peers, and alcohol use. Regarding psychological characteristics of adolescents, the authors found that higher levels of self-esteem and self-confidence were negatively related to early intercourse.

O'Sullivan (1999) studied mother-daughter communication about sexuality within minority families. As part of an HIV-intervention study, 110 Hispanic adolescent girls (ages 13-18) and their mothers completed questionnaires assessing communication patterns and adolescent sexual behavior. It was found that the quality of general communication between mothers and daughters was more useful in predicting the onset of sexual experiences than specific communication about sex.

A report issued by the National Center on Addiction and Substance Abuse (1999) recommends that substance use and sex should be addressed together in prevention programs. The report is based on analyses of national data sets encompassing more than 34,000 teenagers, as well as reviews of hundreds of relevant articles. The report found that there is a strong connection between substance use and sexual behavior among teenagers. For example, 63 percent of teens who use alcohol and 70 percent of those who drink frequently have had sex, compared with 26 percent of teens who have never used alcohol. Similarly, 72 percent of teens who use drugs and 81 percent who are heavy drug users have had sex, compared with 36 percent of those who have never used drugs. This study suggests that teens who engage in one high-risk behavior are more likely to engage in another, thus it is recommended that schools create comprehensive education programs addressing the relationship between substance abuse and sex (Alcoholism and Drug Abuse Weekly, December, 1999).

Because of the overlap in fundamental information presented to students through sex education and drug education, it seems likely that sex education would have an effect on students' drug behavior and drug education would have an effect on students' sexual behavior. The present study asks whether Project CARE (Community Awareness & Relationship Education), a relationship education program aimed at postponing sexual intercourse had any impact on drug use among students.

Project CARE is a three-year intervention using an abstinence based sexuality education curriculum. Changes in student attitudes and behavior were evaluated during the first year of the program (Goldfarb, Donnelly, Duncan, Young, Eadie & Castiglia, 1999). Significant differences were found which, although small, can be attributed to change in the intervention group. In terms of behavioral outcome, intervention participants reported a significantly lower incidence of sexual intercourse than did the control group and were less likely to believe that they would have sexual intercourse before finishing high school. Given the commonalties between the issues addressed in this curriculum, such as decision-making, self-esteem and communications skills, and those addressed in many drug education curricula, it seemed reasonable to expect that this intervention would have some impact on drug attitudes or behaviors.

METHOD

Procedure

Project CARE is a community test of the effectiveness of an abstinence sexuality education program, the focus of which is to delay the initiation of sexual activity among young adolescents. The intervention consists of the *Sex Can Wait* upper elementary (Young & Young, 1997b) and middle school curricula (Core-Gebhart, Hart, & Young, 1997), and the *Abstinence: Pick and Choose Activities* book (Young & Young, 1997a) implemented respectively in grades six, seven and eight. This intervention clearly promotes sexual abstinence while providing factual information regarding sexual behavior, adolescent pregnancy, child rearing and their consequences. In addition, the curricula emphasize skills for goal setting, decision-making, effective communication, and self-esteem enhancement. (Core-Gebhart, Hart, & Young, 1997). Unlike many abstinence-only curricula that use fear and moral preaching to get their message across, *Sex Can Wait* portrays sexuality as a healthy, normal and very positive aspect of human life. Within this context, it promotes abstinence from sexual intercourse as the best choice young adolescents can make to maintain their physical, psychological and social health.

The *Sex Can Wait* series has three age-appropriate curricula: upper elementary, middle school, and high school. The curriculum is modified for each grade level, in order to retain the interest level of students. In this study, the upper elementary and middle school components are utilized in grades six and seven, respectively, and the related program, *Abstinence: Pick and Choose Activities*, is being implemented in the eighth grade. The curricula are driven by the principles of social learning theory, including emphasis on susceptibility, self-efficacy, social support, peer norms, and skill building. One of the strengths of the curriculum is its user-friendly format, which includes a lesson overview, behavioral objectives, a time estimate, and teacher instructions. It also includes handouts, parent-child homework assignments, guidelines for teacher implementations, and information for parents on how to communicate with their children about sex.

As part of Project CARE, intervention group teachers participate in a three-day intensive training workshop on the implementation of the *Sex Can Wait* curriculum. Teachers then administer a pre-test questionnaire to students, designed to measure sexuality-related attitudes, behaviors, behavioral intentions, and knowledge. Students then receive twenty-three sessions of the *Sex Can Wait* curriculum throughout the school year. At the end of the year, teachers administer a post-test questionnaire.

Teachers in the control/comparison group schools received a one-day workshop in the administration of the questionnaire and then administered a pre-test to their students before beginning whatever health education curricula were already in place. Teachers administered a post-test upon completion of their curricula.

Measures

The evaluation instrument was a 47-item questionnaire. Five questions assessed demographic information, nine related to self-esteem, fifteen to attitudes regarding sexuality, four to sexuality knowledge, nine to behaviors, and five to behavioral intentions. All self-esteem and some attitude questions were assessed via a four-point Likert-type scale. The self-esteem scale had a coefficient alpha internal consistency reliability of .75. Items such as "My parents are proud of the kind of person I am" were followed by a four-point response ranging from "strongly agree" to "strongly disagree". Attitude questions were true/false, such as "I feel that there is nothing wrong with a person my age smoking cigarettes." Behaviors and behavioral intentions were assessed using Likert-type scales or yes-no responses. This instrument was constructed for the project and items came from Kirby's national evaluation of sexuality education programs (1985) and from the original evaluation project for *Sex Can Wait* (Denny, Young, & Spear, 1995).

Participants

Evaluation was based on a classic pre-post control group design. The curriculum was implemented in classes at three intervention schools in an urban setting. Three control schools continued their usual curricular offerings but were prevented from adopting the intervention curricula. Students in sixth, seventh, and eighth grades at each school completed (subject to passive parental consent) a pretest survey which was repeated as a post-test at the end of the academic year. Posttest survey results from 839 students were examined, with 426 students (227 females and199 males) from intervention schools and 413 students (211 females and 202 males) from control schools.

Fewer than 2% of the students did not participate due to lack of parental consent. Passive parental consent was approved by the Montclair State University Internal Review Board Human Subjects Committee as an appropriate means of consent for this population. All parents received a passive consent form discussing the questionnaire and program. They were instructed to return the form with their signature if they did not approve of their child's participation in the program or in completing the questionnaire.

Data analysis was nonparametric in nature, involving no *a priori* assumptions about the underlying distributional characteristics of the data. Analysis included pre to post comparisons for the total subject population and for the intervention and control groups separately, as well as posttest only comparisons between the intervention and control groups.

RESULTS AND DISCUSSION

While pretest comparisons had shown the intervention and control groups to be comparable, this was not true for all variables for the students completing the posttest. At posttest, there was a small but significant difference ($p < .02$) in the age distribution of the two groups, such that fewer of the control subjects fell within the oldest age group and highest grade in school. Given that likelihood of drug use increases with age among adolescents, this creates a small bias against finding a positive impact of the intervention on drug use behavior.

TABLE 1 Prevalence of Drug Use for Past Month by Drug and Group

	Controls (n= 413)		Intervention (n= 426)		
	%	n	%	n	*p*
Alcohol	12.1	50	13.8	58	.474
Tobacco	5.8	24	9.4	40	.068
Marijuana	6.8	28	4.0	17	.091
Inhalants	5.3	22	3.3	14	.173
Cocaine	3.9	16	3.0	13	.573
Crank	5.1	21	2.3	10	.043
Heroin	3.4	14	2.1	9	.294

It was also found at posttest that a majority (53.7%) of the intervention students lived in a house, while a minority (46.3%) of the control students lived in a house ($p < .04$). This may reflect an unintended socioeconomic bias in the posttest sample resulting from differential attrition from the pretest group. There were no other significant differences between the groups in terms of either demographics or living arrangements. The two groups did not differ significantly in either their frequency of attendance at religious services or their self-assessed degree of religiosity.

Improved self-esteem was one of the objectives of the curriculum and was measured by nine items on the survey. No differences between the groups were found on the students' responses to either the statement that, "I think that I have good personal qualities," "I am comfortable making my own decisions," and "I know what to do in order to reach my life's goals." Nearly 90% of the students had endorsed each of these items at pretest, leaving little room for improvement. The change from pretest to posttest was, in fact, less than 1% on each item.

Small but significant ($p < .02$) differences between the two groups were found on three self-esteem items. Students exposed to the curriculum were more likely than control students to agree or strongly agree that, "I value respect for myself and others" and that "I can see myself graduating from high school". On the other hand, control students were more likely to agree or strongly agree that, "I feel good about myself on a regular basis". In each case more than 80% of both groups agreed with the item and the difference was between 3.5 and 6.9% in the proportion agreeing.

No group differences were found on items intended to measure social support derived from family and friends, such as "I am an important person to my family", and "My parent(s)/guardian(s) believe that I will be a success in the future." Again this was largely due to the fact that most students responded favorably on these items at pretest.

As can be seen in Table 1, the intervention and control students did not differ significantly in the proportions reporting past month use of alcohol, tobacco, marijuana, inhalants, cocaine/crack, or heroin. Controls were significantly more likely to use methamphetamine (commonly known as "crank" or "ice"). Self-reported use of crank or ice had increased significantly in both groups from pre to post -- reported by 1.4% of controls at pretest and 5.1% a year later ($p < .002$), and by 0.5% and 2.3% of intervention subjects over the same period ($p < .02$). Use of cocaine or crack also increased significantly in the control group, from 1.6% to 3.9% ($p < .05$) but not in the intervention group, which increased only slightly, from 1.5% to 1.9%.

The results of this evaluation do not suggest that a relationship based sex education curricula had meaningful spillover effects on drug-related behaviors. In general the results are consistent with the view proposed by Kirby et al. (1994) that effective health behavior change will result only from narrowly focused curricula with specific behavioral goals.

A difference was found, however, in the prevalence of use of amphetamines -- "crank" or "ice." Furthermore, cocaine use increased from pre to post among controls but not intervention subjects These differences could, of course, be due solely to chance. It seems worthwhile, however, to consider the possibility that the small minority of students at risk for use of these drugs might be a group particularly in need of assistance with problems of relationships, communications and self-esteem. Further examination of this issue seems to be indicated.

REFERENCES

Alan Guttmacher Institute web site: www.agi-usa-org. 1999.

Alcoholism and Drug Abuse Weekly, December 1999. Report: Prevention Curricula Should Address Drug Use and Sex, *11*(48), 5-6.

Black, D. & Tobler, N. (1998) Peer Helping/Involvement: An Efficacious Way to Meet the Challenge of Reducing Alcohol, Tobacco, and other Drug Use Among Youth? *Journal of School Health*, 68(3), 87-94.

Core-Gebhart, P., Hart, S. and Young, M. (1997). *Sex Can Wait – a curriculum for middle school classrooms - OAPP Approved Edition*. Santa Cruz, CA: ETR Associates.

Denny, G., Young, M., & Spear, C. (1995). "Sex Can Wait" Training and Evaluation Project; Student Survey Booklet." Unpublished Test Instrument.

Donnelly, J., Goldfarb, E., Duncan, D., Young, M., Eadie, C., Castiglia, D. (1999). Self-Esteem and Sex Attitudes as Predictors of Sexual Abstinence by Inner-City early Adolescents. *North American Journal of Psychology*, *1*(2): 205-212.

Drug Strategies. (1996). *Making the Grade: A guide to school drug prevention programs*. Washington, DC: Author.

Ferguson, S. (1998). Peer Counseling in a Culturally Specific Adolescent Pregnancy Program, *Journal of Health Care for the Poor and Underserved*, *9*(3), 322-341.

Hofler, M., & Lieb, R. (1999). Covariates of Cannabis use Progression in a Representative Population Sample of Adolescents: A Prospective Examination of Vulnerability and Risk Factors. *Addiction*, *94*(11): 1679-1695.

Humphries, S. (1999). Improving Peer Mediation Programs: Student Experiences and Suggestions, *Professional School Counseling*, *3*(1), 13-21.

Johnston, L.D., O'Malley, P.M. & Bachman, J.G. (2000). *Monitoring the Future national survey results on drug use, 1995-1999. Volume 1: Secondary school students*. Bethesda, MD: National Institute on Drug Abuse.

Kawabata, T., Cross, D., Nishioka, N., Shimai, S. (1999). Relationship between Self-Esteem and Smoking Behavior Among Japanese Early Adolescents: Initial Results from a Three-Year Study. *Journal of School Health*, *69(7):* 280-285.

Kirby, D., Short, L., Collins, J., Rugg, D., Kolbe, L., Howard, M., Miller, B., Sonenstein, F., and Zabin, L. S. (1994). School-based Programs to Reduce Sexual Risk Behaviors: A Review of Effectiveness. *Public Health Reports*, *109*(3), 339-360.

Kirby, D. (1985) "Sexuality Education : An Evaluation of Programs and their Effects." ETR Associates, Santa Cruz,CA.

Levy, D.T., & Stewart, K. "Costs of Underage Drinking." Report prepared for the US Dept. of Justice, Office of Juvenile Delinquency Prevention, Pacific Institute for Research Evaluation, July, 1999.

Mayer, O. (1999). Substance Abuse Prevention Takes to the Classroom. *State Legislatures*, *25*, 24-30.

National Center on Addiction and Substance Abuse. (1999). *Dangerous Liaisons: Substance abuse and sex*. New York, NY: Columbia University.

O'Sullivan, L, Jaramillo, B., Moreau, D., Meyer-Bahlburg, H. (1999). Mother-daughter communication about sexuality in a clinical sample of Hispanic adolescent girls. *Hispanic Journal of Behavioral Sciences, 21*(4): 447-470.

Reasoner, R. (2000). *Self-Esteem and Youth: What Research Has to Say About It.* International Council for Self-Esteem, Port Ludlow, WA.

Whitbeck, L., Yoder, K., Hoyt, D., Conger, R. (1999). Early adolescent sexual activity: A development study. *Journal of Marriage & the Family, 61*(4): 934-947.

Young, M. & Young, T. (1997a). *Abstinence Pick and Choose Activity Book.* ETR Associates, Santa Cruz.

Young, M., & Young, T. (1997b). *Sex Can Wait: An Abstinence Based Sexuality Education Program for Upper Elementary Classrooms.* ETR Associates, Santa Cruz.

Note: This project was made possible with the funding through the U.S. Dept. of Health & Human Services, Office of Adolescent Pregnancy Programs.

Attribution and Reciprocity in International Relations: The Attribution Reciprocity Model

David Daniel Bogumil
Wright State University

Understanding dyadic interaction in relationships represents an enduring challenge to reciprocity research. The application of many paradigms has been used to discern patterns of reciprocity in dyadic interaction: interpersonal, intergroup, inter-organizational, and international interaction. The advancement of dyadic relations and reciprocity research requires both theoretical and methodological innovations to further our understanding of the elusive nature of patterns of reciprocity (Baldwin, 1998; Druckman, 1998; Larson, 1998). Stern and Druckman (2000) have suggested that research in international relations requires the exegesis of conceptual issues of the nature of actions pertinent to the process of reciprocity. This paper presents an innovation in reciprocity research by the introduction of the Attribution Reciprocity (AR) model. The AR model presents a theoretical catalyst for development of dyadic interaction and reciprocity research. This article examines the application of attribution theory to the dyadic interactions of nation states.

The analysis of reciprocal actions taken by nations can benefit from an understanding of the attribution process. In past studies, a nation's propensity to engage in reciprocity and the type of reciprocal interaction has subsumed the multidimensional complexity of attribution states within a continuum of conflict and cooperation. Numerous scholars have studied and investigated various conceptualizations of the cooperative and conflictive nature of an action or event (Azar, 1980; Azar & Havener, 1976; Calhoun, 1971; Druckman, 1998; Goldstein, 1992; Howell, 1979; McClelland, 1978; McClelland, 1983; Merritt & Zinnes, 1988; Richman, 1975; Stern & Druckman, 2000; Vincent, 1983). However, past research has focused on a continuum of conflict and cooperation. Recently, Stern and Druckman (2000) have suggested that the advancement of reciprocity research in international relations requires the clarification of conceptual issues of the codification of actions. A significant development in the analytic precision of reciprocity research is possible with the understanding of the attribution states of international interactants. The application of attribution theory can enhance our understanding of the actions of nations through a multidimensional

---

Author info: Correspondence should be sent to: David Daniel Bogumil, Ph.D., Dept. of Sociology and Anthropology, Wright State University, Millett Hall, 3640 Colonel Glenn Highway, Dayton, OH 45435. david.bogumil@wright.edu

North American Journal of Psychology, 2001, Vol. 3, No. 3, 463-480.

© NAJP

evaluation of reciprocity and the nature of conflict and cooperation based on attribution states.

This paper proposes the application of attribution theory to the actions of nation states. The attribution reciprocity (AR) model developed in this paper provides a unique opportunity to clarify some persistent conceptual issues in the study of reciprocity in international relations (Weiner, 1972, 1979, 1992; 1995; 1998; Weiner & Litman-Adizes, 1980; Weiner, Nierenberg, & Goldstein, 1976).

Reciprocity between nation states has been examined in terms of the levels of cooperation and conflict (Azar, 1980; Bogumil, 1993; Freeman & Goldstein, 1989; Patchen, 1988; Patchen & Bogumil, 1995, 1997; Stern & Druckman, 2000). The utility of a uni-dimensonal cooperation conflict scale has been used to rate the implicit actions of the leadership of nation states and their public statements on the character of international events (Holsti & Rosenau, 1988; Larson, 1998; Parks & Komorita, 1998; Patchen, 1998). This process of understanding the action or event's cooperative and conflictive valence and the effect of that interpretation of events on our capacity to detect the presence of patterns of reciprocity is basically an exercise in attribution analysis. Therefore, one's understanding of the process of reciprocity can benefit from a theoretical perspective that improves one's understanding of the key elements of interaction. Attribution theory can illuminate the attribution states of the interactants and those states' effect on the conflictive and cooperative actions of dyadic interactants.

Toward A New Taxonomy of Reciprocity: The Attribution Reciprocity (AR) Model

The case for reciprocity is clearly identified within the literature. However, the major schism in the literature exists not with the persistence of reciprocity, but with whether one can improve the conceptual understanding of reciprocity to account for the sparse and contextually constrained identification of incidents of reciprocity (Stern & Druckman, 2000). Moreover, the desiderata is a conceptual understanding that will yield a taxonomy and operationalization of variables that will assist in revealing patterns of reciprocity. These presently invisible patterns of reciprocity can be discerned in the data using attribution theory as the basis for the taxonomy and operationalization of key variables. The identification of these variables is critical to understanding the patterns of reciprocity. The evidence of reciprocity in past studies reveals modest findings due to the constraints of uni-dimensional scales—the scales used to measure and score actions on a continuum of conflictful to cooperative behavior (Azar, 1980; Azar & Havener, 1976; Calhoun, 1971; Druckman, 1998; Goldstein, 1992;

McClelland, 1978; McClelland, 1983; Merritt & Zinnes, 1988; Richman, 1975; Stern & Druckman, 2000; Vincent, 1983).

Historically, reciprocity has been commonly found in international relations. In general, the occurrence of reciprocity usually takes the form of the reciprocation of either positive or negative actions. Of course, the reciprocation of positive actions more than the reciprocation of negative actions is indicative of cooperative behavior. Essentially, the perception of a positive action can affect whether the recipient of that action ultimately defines it as positive. For example, Larson's (1988) study of reciprocity in superpower relations found that positive actions or favors will be reciprocated only after a state's evaluation of the actual and perceived value of a response.

Larson (1988) states: "The motives perceived to underlay another state's proposals affect whether it is viewed as a concession and shape the formulation of response" (p. 298). There are many theoretical explanations for reciprocity. However, a majority of these explanations portray self-interest as a motivation for reciprocation (Macy, 1989). Moreover, the reciprocation of an action may determine a level of compliance and in some cases an acceptance of a condition of threat. The attributions manifest in the actions of either member of a dyad requires a theory that addresses more than the positive or negative dimension of an action or event. The identification of an attribution of an event or action as a threat recognizes that the respondent and the actor are engaged in a more complex interaction.

A taxonomy that uses attribution theory provides both for the clarification of the complexity of the existing models of simple, trend, and comparative reciprocity and a dictate for the reanalysis of these established reciprocity models. Furthermore, I suggest that existing models of reciprocity require a metamorphous into new hybrid models based on the foundations of simple, trend, and comparative models using attribution theory. The modest evidence for reciprocity in all models can be artifact of uni-dimensional scales and their assorted weighting schemes analysis (Azar, 1980; Azar & Havener, 1976; Calhoun, 1971; Druckman, 1998; Goldstein, 1992; Howell, 1979; McClelland, 1978; McClelland, 1983; Merritt & Zinnes, 1988; Richman, 1975; Stern & Druckman, 2000; Vincent, 1983).

I identify my proposed model as the Attribution Reciprocity (AR) model. The AR model can reveal the presently imperceptible complexity of actions of reciprocity that the simple, trend, and comparative reciprocity models are based upon (i.e., the unidimensional scales used in the Conflict and Peace Data Bank—COPDAB (Azar, 1982) and World Events Interaction Survey—WEIS (McClelland, 1978). The current evidence of reciprocity is conceptually constrained due to the heretoforeunexamined importance and recognition of the attribution process in all levels of dyadic interaction: interpersonal, intergroup, interorganizational, and international relations. To that end, the exposition of the pertinent theoretical evidence and empirical research based in the reciprocity literature illuminates the pragmatic value of the Attribution Reciprocity (AR) model.

Interpersonal Relations

The theoretical cross-fertilization of attribution and reciprocity paradigms for the purpose of the advancement of international relations inquiry requires the integration of pertinent research on the largely separate traditions of attribution and reciprocity. The study of attribution on an interpersonal level (Heider, 1944, 1958; Kelley, 1972a, 1972b; Weiner, 1971, 1979) is well documented in the literature.

The study of attribution in international relations is a new area of inquiry; however, the precepts of the attribution process have a long ubiquitous presence in the reciprocity literature. Historically, Gouldner's (1960) norm of reciprocity states that individuals are socialized into exchange relations. Briefly, people should help those who help them and not hurt those who help them ("quid-pro-quo").

The study of attribution in interpersonal and intergroup relations based on Weiner (1979) suggests a causal chain and a temporal duration of the chain of events that is central to the issue of the attribution process. Reciprocity is an artifact of a more conceptually complex process and may reveal a broader array of interaction patterns of reciprocity (simple, comparative, and trend models). However, the identification of the interactants relevant to the attribution process may also be accomplished by identifying their actions within Weiner's model.

Weiner et al.'s (1971) expression of "locus of causality" offers a critical restatement and modification to the analysis of attribution of Rotter's earlier work on "locus of control" (Rotter, 1966). Weiner presented (1979, 1986) a three dimensional taxonomy of causality of the attribution process. The first dimension identifies an attribution's locus as internal or external. The second dimension identifies an attribution as stable or unstable. The third identifies an attribution as controllable or uncontrollable.

The influence of prior actions of an individual and/or a group will influence the attribution process of an individual. Kelley's (1972a) work asserts that individual motivation for the exercising of control affects the process of attribution. Kelley explains that the temporal events of covariation between effect and response (i.e., "law of effect") accumulate *a priori* to an interaction event and serve as indicators in the attribution process. Sprecher (1998) maintains that patterns of reciprocal liking have a cumulative effect that influences attraction based on the expressions of intrinsic satisfaction of interactants. Earlier, Kelley (1972a) suggested that the locus of causality and control of these interactions are unique effects. Kelley presented several principles of discounting and augmentation with precepts of facilitation and inhibition to encompass the separation of a multiplicity of effects with the association of any one cause's locus of causality and control. Molm, Peterson and Takahashi (1999) attributed locus of causality and control based on the power and availability of interactants in reciprocal exchanges networks. They found that risk reduction was accomplished by reciprocal interaction with individuals who were less powerful and therefore more predictable in unilateral actions. Anderson, Guerrero, Buller, and Jorgenson (1998) found that immediacy of reciprocal actions may adversely affect the response of the interactants. Paese and Gilin (2000) assert that reciprocity is feigned and self-interest is the motivation of the interactants. Kenny, Mohr, and Levesque (2001) suggest that individuals within relationships will engage in reciprocity even when a variance in the behaviors of partners is present. The acts of reciprocity may sustain a relationship through any series of crises present in the shared maintenance of long-term relationships.

Group and Organizational Relations

The role of reciprocity as a financial, social and legal bond, positions the concept of reciprocity at the center of social organization (Fausto, 1999). Martinko (1995), McDonald (1995) and Weiner (1995) maintain that motivation and response of group and organization members and intergroup and interorganizational entities can be evaluated within a rational attribution taxonomy.

The cognitive tradition of Heider's (1958) work on interpersonal relations was expanded by research on group dynamics and individual behavior (Cartwright and Zander, 1960). Zander's (1971) compendium of experimental work on group motivation and beliefs reveals that group cohesion is a function of a synchronicity of member and group desires and goals. The maintenance of group identity requires member competence. The member's competence and the group's success are critically intertwined in the perception of the member, the group, and the social pressure of other groups on the group (Zander, 1971, p. 200).

Raven, Schwarzwald, and Koslowsky's (1998) evaluation of power in reciprocity maintains that a subordinate member of an organization independently may assess the informational and expert power of a supervisor based on the subordinate's perception of the locus of causality (external or internal) of the supervisor's actions. Likewise, the inference of members of in-groups and out groups form schema and subsequently attribute behaviors of "the other" group as common and salient characteristics of that group (Heiss, 1981; Kanouse, 1972; Kelley, 1972b; Singer, 1981). This synergy of boundary maintenance and cognitive consistency functions sustains group identity. Attributions common to group members have a greater impact on groups that are interdependent than groups that are independent (Lu, 1997; Zaccaro, Peterson, & Walker, 1987). The salience of attributions made by groups within a context of a shared symbolic environment is an artifact of the interactions among the groups and their environment (Lippitt, Polansky, & Rosen, 1952; Shavitt & Fazio, 1991). Rotter (1966) regards the internal dimension as contingent on one's "own relatively permanent characteristics" and the external dimension as "the great complexity of forces surrounding" the individual. The interaction and negotiation of a group's perception of internal and the external dimensions of attribution and the other dimensions of the attribution process represents the group's valance. The group's judgement affects their behavior. The valance of internal and external attribution foments group members' cognitive judgment and their perception of "other" groups (Heiss, 1981).

Buunk, Doosje, Jans, and Hopstaken (1993) found that patterns of reciprocity were aligned with the interests of the norms and values of the organization. They stated that subordinates shared perceptions of patterns of reciprocity. Subordinates perceptions of the reciprocity of colleagues were expressed as balanced interactions juxtaposed to their expressions of imbalanced reciprocal interactions with supervisors.

Interorganizational Relations

McDonald (1995) explains the attribution process as a function of interdependent group social identity (Bogumil, 1999; Brewer, 1985; Brewer & Schneider, 1990; Brown & Wooten-Millward, 1993; Deschamps, 1973-1974; Doise et al, 1988; Tajfel, 1973) within an organizational environment. The members of groups maintain their schema by rational attribution of causality of their status. Furthermore, their status reifies their casual schema that informs their rational attributions regarding their group, the other group, and the environment (Branscombe, Schmitt, & Harvey, 1999; Oetting, 1999; Weiner, 1998). Laumann and Marsden's (1982) study of reciprocity in inter-organizational settings found that in a highly competitive environment, autonomous organizations would act with regard to self-interest.

In the case of "autonomous organizations engaged in voluntarily coordinated pursuit of collective goals," there exists an expectation of reciprocity based on shared values (Laumann & Marsden, 1982, p. 344). Laumann and Marsden suggest these similarities in dyads serve to increase mutuality. They illuminate an approach that emphasizes the basis for future interorganizational studies examining reciprocity--namely, to employ the theory based in interpersonal relations to structure further inquires.

Michener, Vaske, Schleifer, Plazewski and Chapman (1975) found that reciprocity between bargainers was affected by a bargainer's concession rate. The more powerful party in negotiation viewed a greater frequency of concession by the "other" party as the less incentive to reciprocate. Eyuboglu and Buja (1993) maintain that inter-organizational negotiations exhibit reciprocity in the process of contention. They suggest that organizations engaged in bargaining commit to patterns of expected knowledge of another organization and misperception of the other creates contention. However, this unbridled contention between negotiating parties may only serve to complicate the development of a shared basis of facts. A tactic of making a "threat" has no meaning when it cannot be perceived amongst many other unintentional actions that foster the same perception of "threat." This unintentional outcome limits proper exploitation of the utility of contention (i.e., making a "threat").

Cialdini (1996) and Settoon, Bennett, and Liden (1997) assert that the embedded patterns of attribution in the perception of other groups (sections, divisions, and/or departments) promotes an organizational environment susceptible to influence systems that can negatively affect an organization's reciprocal relations (i.e., a corporate culture).

International Relations

Attribution theory has been examined extensively at the interpersonal level of dyadic relations. Likewise, the study of models of reciprocity has occurred within the domain of interpersonal relations (Downie & Robbins, 1998; Kenny, Mohr, & Levesque, 2001; Lu, 1997; Thibaut & Kelley, 1967; see Bogumil, 1996, 2001 for an extensive review).

Furthermore, there is a substantial literature within the study of international relations that has dealt with cooperation, conflict and reciprocity (Bogumil, 1993; Dixon, 1986; Druckman & Harris, 1990; Freeman & Goldstein, 1989; Goldstein & Freeman, 1990; Greffenius, 1990; Larson, 1988; Leng, 1993; Patchen, 1988; Patchen, 1991; Patchen & Bogumil, 1995, 1997; Richardson, Kegley, & Agnew, 1981; Sigler, 1972; Starr, 1978; Stern & Druckman, 2000; Stoll & McAndrews, 1986; Van Wyk & Radloff, 1993).

There is modest evidence that national leaders generally reciprocate the actions of others under various conditions (Bogumil, 1993; Holsti & Rosenau, 1988; Patchen, 1988: 262-263). This tendency toward reciprocity in international relations has been empirically found in interactions between the USSR and the US (Dixon, 1986; Druckman & Harris, 1990; Freeman & Goldstein, 1989; Goldstein & Freeman, 1990;

Greffenius, 1990; Larson, 1988; Leng, 1993; Patchen, 1988; Patchen, 1991; Patchen & Bogumil, 1995, 1997; Pruitt, 1969; Richardson, Kegley, & Agnew, 1981; Sigler, 1972; Starr, 1978; Stern & Druckman, 2000; Stoll & McAndrews, 1986).

The study of the attribution as integral to the process of reciprocity is a relatively unexamined area of inquiry. Dyadic relationships have been evaluated using either attribution or reciprocity/social exchange theoretical perspectives from the interpersonal to the international levels of analysis. The nation state provides a change in scale; however, the dynamics are grounded in the international relations (IR) literature and are fundamental to process of reciprocity in the IR literature (Braver, 1975; Bremer, 1992; Larson, 1998; Leng, 1998; Patchen, 1998; Stern & Druckman, 2000). The inferences of international actors (nation states) form schema and subsequently they attribute actions to common characteristics of the actor (Bogumil, 1993; Cialdini, Wosinka, Barrett, Butner, & Gornik-Durnose, 2001; Heiss, 1981; Kanouse, 1972; Kelley, 1972a, 1972b; Singer, 1981; Wosinka, Cialdini, Barrett, & Reykowsi, 2001).

This relationship between boundary maintenance of a nation-states projected identity in international relations requires cognitive consistency in their international policy. The synergy of the maintenance of stability in their policy and the functions of leadership in the implementation of that policy can result in the bureaucratic momentum of group identity (Brady & Kegley, 1977; Dixon, 1986). Janis (1989) identified this potentially deleterious "mindguarding" process as one basis for the process of Groupthink. One may understand this as volatile state of the presence of inertia in decision models.

Reciprocity research subsumes collective identity and leadership action as tandem and coupled elements in the international interaction process (Brewer, 1985; Brewer & Schneider, 1990; Brown & Wooten-Millward, 1993; Cartwright & Zander, 1960; Dalby & Wertman, 1971; de Reuck, 1988; Doise, Turner, Rabbie, & Horwitz, 1988; Janis, 1989; Larson, 1998; Leng, 1998; Patchen, 1998; Ward, 1982; Zander, 1971). Furthermore, the presence of international organizations can foster attributions common to nation state group members and may have a greater impact on groups that are interdependent than groups that are independent (Cusack & Eberwein, 1982; Dixon, 1990; Druckman, Broome & Korper, 1988; Ekhardt & Azar, 1978; Ward & Widamaier, 1982; Zaccaro, Peterson, & Walker, 1987). Many factors may affect the occurrence of reciprocity between nations. These factors include cultural differences (Hwang, 1987), differences in the values related to human rights (Van Wyk & Radloff, 1993), economic strength or unequal power (Leng, 1993; Patchen, 1988), domestic policy beliefs about foreign policy initiatives (Holsti & Rosenau, 1988), international negotiator communication styles (Druckman, Broome & Korper, 1988) and elections cycles (Gaubatz, 1991). The reciprocal response of a nation has been attributed to a golden rule of international politics set forth by Ward (1981). Ward's modified golden rule's basic premise is "to do unto others as they have recently done unto you." Simply stated, the reciprocation of a past action of another finds modest support in a wide variety of analyses (Bogumil, 1993; Patchen, 1988; Patchen & Bogumil, 1995, 1997, Stern & Druckman, 2000). Stern and Druckman (2000) have identified the systemic nature of reciprocity research based in existing conceptualizations of the actions of international actors or nation states. The application of Weiner's model of attribution theory to the reconceptualization of reciprocity research in international relations can provide important innovation to the study of international relations.

The AR Model: Transition to an Attribution Taxonomy

Research using the Attribution Reciprocity (AR) model stipulates a required articulation of the manner of implementation or transition to the AR model. The transition to the AR model can be accomplished by recoding the existing narratives using the attribution taxonomy. For example, the AR model's coding scheme would require transformation of an attitude object (i.e., the narrative statement) to be evaluated in terms of the three dimensions. The attitude object (i.e., the narrative statement) would receive a value on the following dimensions: internal or external, stable or unstable, and controllable or uncontrollable. The coding of the actions relative to the states of attribution and the positive or negative valence of the action will permit a more precise analysis of patterns of reciprocity. The coding process would necessitate a research protocol that included an evaluation of the validity and reliability of the judges/coders.

A recoding of existing scale values (Azar, 1980; Azar & Haverner, 1976; Calhoun, 1971; Vincent, 1983) is problematic. That process may undermine the capturing of the "true" multidimensional nature of an action/event. For example, an attempt to employ that type of coding scheme would require transformation of an attitude object word such as "threat" to be evaluated in terms of the three attribution state dimensions. The expression "threat" would receive a value on the following dimensions: internal or external, stable or unstable, and controllable or uncontrollable. The transformation of single descriptor without the context of descriptors into a three dimensional expression of attribution states raises issues of validity.

The Attribution Reciprocity (AR) model's dimensions are expressed as dichotomous variables—mutually exclusive attribution states (e.g., internal or external, stable or unstable, and controllable or uncontrollable). However, one may extend these dichotomous dimensional characteristics of attribution to expressions of continuous dimensions of attribution states (Weiner, 1986). This suggestion would be similar in method to Bales (1979) SYstem for Multiple Level Observation of Groups (SYMLOG)—a three-dimensional cube design with spatial characteristics of proximity of personality. The methodological extension beyond the dichotomous categories has implications for the present theoretical dimensionality of the AR model. Moreover, these two intertwined aspects of method and theory intersect in a potentially precarious and problematic manner.

The following is a proposed and limited AR model treatment of the recent case of US relations with China. The scope of this paper is the theoretical development of the AR model for the study of reciprocity in international relations. This paper is not an empirical examination that mandates rigorous statistical analyses of event coding procedures and event patterns of reciprocity. Therefore, this example of the AR model must be evaluated in a circumspect fashion given they lack the empirical evaluation of the cumulative reciprocal events. The empirical evaluation of events is imperative as context for understanding the dyadic interactions of nation-states. The absence of a context based on an empirical examination can only be replaced with an anecdotal historical context. Furthermore, this exposition of the principles of the AR model serves to merely identify the attribution states of the nation-states. The inference of patterns of reciprocity: simple, trend, or comparative are authentic and conclusive, subject to empirical analysis (Bogumil, 1993).

The following events represent provide an account of the recent case of the Chinese government's interception and forced landing of a US reconnaissance plane. This example's purpose is to present an identification of the attribution states proposed in the AR model. This example identifies the attribution states of the US and China dyad during this episode. The AR model presentation in this example identifies the dichotomous categories of attribution states: external or internal, stable or unstable, and controllable or uncontrollable. The reciprocity and the attribution states identified and assigned to the actions of the US and China must be understood in their historical context.

> US-China episode: A US Navy EP-3 surveillance aircraft with a twenty-four member crew collided with a Chinese fighter on April 1, 2001.
>
> Near China's coastline a US Reconnaissance plane encounters a Chinese fighter shadowing the US plane. Contact between the planes occurs and the Chinese fighter jet crashes into the sea and the pilot Wang Wei dies in the crash.

Identification of Attribution States
US Perception of US: External, Stable, Uncontrollable
US Perception China: Internal, Stable, and Controllable
China Perception of US: External, Stable, and Uncontrollable
China perception of China: Internal, Stable, Uncontrollable

President Bush states the US plane was over International waters and states the US will not offer and apology. China states that the actions of the US plane and its crew were premeditated.

Identification of Attribution States
US Perception of US: External, Stable, Controllable
US Perception China: Internal, Stable, and Controllable
China Perception of US: Internal, Stable, and Controllable
China perception of China: External, Unstable, Uncontrollable

U.S. Ambassador to China Joseph Prueher writes a letter to Chinese Foreign Minister Tang Jiaxuan. The letter states that the plane's crew had no option but to land in China given the crippled condition of the plane. The US letter states its regret for the death of the Chinese pilot and the US plane's unattended violation of China airspace due to the crippled status of the plane.

Identification of Attribution States
US Perception of US: External, Stable, Controllable
US Perception China: Internal, Stable, and Controllable
China Perception of US: Internal, Unstable, and Controllable
China perception of China: External, Stable, Controllable

The Chinese government announces the release of the US plane's crew.
Identification of Attribution States
US Perception of US: Internal, Stable, Controllable
US Perception China: Internal, Stable, and Controllable
China Perception of US: Internal, Stable, and Controllable
China perception of China: Internal, Stable, Controllable

Secretary of State Powell states that China is easing the pursuit of U.S. jets.
Identification of Attribution States
US Perception of US: Internal, Stable, Controllable
US Perception China: Internal, Stable, and Controllable
China Perception of US: Internal, Stable, and Controllable
China perception of China: Internal, Stable, Controllable

Secretary of State Powell is upbeat about a warmer relationship.
Identification of Attribution States
US Perception of US: Internal, Stable, Controllable
US Perception China: Internal, Stable, and Controllable

China Perception of US:	Internal, Stable, and Controllable
China perception of China:	Internal, Stable, Controllable

This example of identification occurs within the wider context of the history of US-China relations. At this stage, both China and the US are reciprocating action across attribution states in preparation for the MFN agreement. These discernable attribution state patterns of reciprocity are similar to periodicity waves. The suggestion of simple reciprocity can be inferred by the reciprocation of stability in attribution states. However, there are multiple AR models operating simultaneously and they enrich the context and must be brought into account. There are many examples of multiple wave manifestation of the AR model in US-China relations in the presence of various international events. One is the allied efforts of the US and China during World War II in the Pacific theater (e.g., the Flying Tigers—a covert private/public enterprise).

The identification of a wave or the intersection of waves in patterns of reciprocity using the AR model is a rigorous process. The review of only a few of the events in a wave reveals the level of complexity of such an analysis. Moreover, the demands of relevant context would also require the evaluation of ongoing policies regarding China's Most Favored Nation (MFN) status, China human rights violations (i.e., religious groups, quasi-religious movements and activists—Falun Gong and dissonant scholars) and the status of Tibet.

Conclusion

The future of reciprocity and attribution research can benefit from the Attribution Reciprocity (AR) model's inclusion and identification of the attribution states that previous scholarship had subsumed on a continuum of conflict and cooperation. Furthermore, the comparative evaluation of previous research findings on reciprocity and attribution will increase our understanding of widely identified patterns of reciprocity within the simple, trend, and comparative models (Bogumil, 1993; Bueno De Mesquita, 1990; Dacey & Pendegraft, 1988; Druckman, 1998; Leng, 1998; Patchen, 1998; Patchen & Bogumil, 1995, 1997; Stern & Druckman, 2000).

The promise of these previously studied patterns of reciprocity may be more fully realized in future research with the implementation of the Attribution Reciprocity (AR) model. Attribution theory provides a unique opportunity to bring clarification to fundamental theoretical, methodological, and conceptual issues confronting the study of reciprocity. The extension of the application of the Attribution Reciprocity (AR) model to the remaining levels of dyadic interaction: interpersonal, intergroup, and interorganizational relations will permit an evaluation of the efficacy of the AR model.

REFERENCES

Anderson, P.A., Guerrero, L. K., Buller, D. B., & Jorgensen, P. F. (1998). An empirical comparison of three theories of nonverbal immediacy exchange. *Human Communication Research, 24(4)*, 501-535.

Azar, E. (1980). The Conflict and Peace Data Bank (COPDAB) Project. *Journal of Conflict Resolution, 24*, 143-152.

Azar, E., & Havener, T. H. (1976). Discontinuities in the symbolic environment: A problem in scaling. *International Interactions, 2*, 231-246.

Baldwin, D. A. (1998). Exchange theory in international negotiations. *International Negotiation, 3*(2), 139-149.

Bales, R. F. (1979). *SYMLOG: A system for multiple level observation of groups*. New York: The Free Press: A Division of Macmillan Publishing Co., Inc.

Bogumil, D. D. (1993). Models of conflict and cooperation: Interaction patterns of international actors [nation-states of the U.S. & U.S.S.R.]. *Dissertation Abstracts International, 55*(03A), 745. (University Microfilms No. 9420786). Doctoral dissertation, Purdue University.

Bogumil, D. D. (1999). *External attribution of stable and uncontrollable dimensions of causality of adolescent substance abuse behavior.* Paper presented at the 1999 Annual Meeting of the American Sociological Association. Section on Social Psychology: Refereed Roundtables Session. Chicago, IL.

Bogumil, D. D. (2001). *Attribution states: Emergency room crisis management.* Paper presented at the 96th Annual Meeting of the American Sociological Association. Regular Session: Health and Well Being. Anaheim, CA.

Brady, L. P., & Kegley Jr., C. W. (1977). Bureaucratic determinants of foreign policy: Some empirical evidence. *International Interactions, 3*, 33-50.

Branscombe, N. R., Schmitt, M. T., & Harvey, R. D. (1999). Perceiving pervasive discrimination among African Americans: Implications for group identification and well-being. *Journal of Personality and Social Psychology, 77*(1), 135-149.

Braver, S. L. (1975). Reciprocity, cohesiveness and cooperation in two-person games. *Psychological Reports, 36*, 371-378.

Bremer, S. A. (1992). Dangerous dyads: Conditions affecting the likelihood of interstate war, 1816-1965. *Journal of Conflict Resolution, 36*(2), 309-341.

Brewer, M. B. (1985). The role of ethnocentrism in intergroup conflict. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (2nd ed., pp. 88-102). Chicago: Nelson-Hall.

Brewer, M. B., & Schneider, S. K. (1990). Social identity and social dilemmas: A double-edged sword. In D. Abrams & M. A. Hogg (Eds.), *Social identity theory: Constructive and critical advances* (pp. 169-184). New York: Springer-Verlag.

Brown, R., & Wooten-Millward, L. (1993). Perceptions of group homogeneity during group formation and change. *Social Cognition, 11*(1), 126-149.

Bueno De Mesquita, B. (1990). Big wars, little wars: Avoiding selection bias. *International Interactions, 16*(3), 159-169.

Buunk, B. P., Doosje, B. J., Jans, L. G. J. M., Hopstaken, L. E. M. (1993). Perceived reciprocity, social support, and stress at work: The role of exchange and communal orientation. *Journal of Personality and Social Psychology, 65*(4), 801-811.

Calhoun, H. L. (1971, April). *The measurement and scaling of event data using the semantic differential.* Paper presented at the Annual Meeting of the Western Political Science Association, Dallas, TX.

Cartwright, D., & Zander, A. (1960). Origins of group dynamics. In D. Cartwright & A. Zander (Eds.), *Group dynamics: Research and theory* (2nd ed., pp. 3-32). New York: Harper & Row, Publishers.

Cialdini, R. B. (1996). Social influence and the triple tumor structure of organizational dishonesty. In D. M. Messick & A. E. Tenbrunsel (Eds.), *Codes of Conduct: Behavioral research into business ethics* (pp. 44- 58). New York: Russell Sage Foundation.

Cialdini, R. B., Wosinka, W., Barrett, D. W., Butner, J. & Gornik-Durnose, M. (2001). The differential impact of two social influence principles on individualists and collectivists in Poland and the United States. In W. Wosinka, R. B. Cialdini, D.W. Barrett, & J. Reykowsi (Eds.), *The practice of social influence in multiple cultures* (pp. 33-50). Mahwah, NJ: Lawrence Erlbaum Associates.

Cusack, T. R., & Eberwein, W. D. (1982). Prelude to war: Incidence, escalation and intervention in international disputes, 1900-1976. *International Interactions, 9*(1), 9-28.

Dacey, R., & Pendegraft, N. (1988). The optimality of Tit-For-Tat. *International Interactions, 15*(1), 45-64.

Dalby, M. T., & Werthman, M. S. (Eds.). (1971). *Bureaucracy in historical perspective.* Glenview, IL: Scott, Foresman and Company.

deReuck, A. (1988). International Relations: A theoretical synthesis. *International Interactions, 14*(4), 321-341.

Deschamps, J. (1973-1974). Attribution, social categorization and intergroup representations. *Bulletin de Psychologie, 27*(13-14), 710-721.

Dixon, W. J. (1986). Reciprocity in United States-USSR relations: Multiple symmetry or issue linkage? *American Journal of Political Science, 30*(2), 421-445.

Dixon, W. J. (1990). Interdependence and cooperation in foreign policy behavior: A comment on Vincent. *International Interactions, 16*(2), 109-114.

Doise, W., Turner, J. C., Rabbie, J.M., & Horwitz, M. (1988). Individual and social identities in intergroup relations. *European Journal of Social Psychology, 18*(2), 99-111.

Downie, M. S., & Robbins, S. B. (1998). Assessing the qualities of social relationships in clinical and nonclinical individuals. *Counseling Psychologist, 26*(3), 466-488.

Druckman, D. (1998). Social exchange theory: Premises and prospects. *International Negotiation, 3*(2), 253-266.

Druckman, D., & Harris, R. (1990). Alternative models of responsiveness in international negotiation. *Journal of Conflict Resolution, 34*(2), 234-251.

Druckman, D., Broome, B. J., & Korper, S. H. (1988). Value differences and conflict resolution. *Journal of Conflict Resolution, 32*(3) 489-510.

Ekhardt, W., & Azar, E. E. (1978). Major world cooperation events, 1945 to 1975. *International Interactions, 5*(2&3), 203-239.

Eyuboglu, N., & Buja, A. (1993). Dynamics of channel negotiations: Contention and reciprocity. Psychology & Marketing, 10(1), 47-65.

Fausto, C. (1999). Of enemies and pets: Warfare and shamanism in Amazonia. *American Ethologist, 26*(4), 933-956.

Freeman, J. R., & Goldstein, J. S. (1989). *US-USSR-Chinese relations: routine, reciprocity, or rational expectations*. Unpublished manuscript.

Gaubatz, K. T. (1991). Election cycles and war. *Journal of Conflict Resolution, 35*(2), 212-244.

Goldstein, J. S. (1992). A conflict-cooperation scale for WEIS events data. *Journal of Conflict Resolution, 36*(2), 365-389.

Goldstein, J. S., & Freeman, J. R. (1990). *Three-way street: Strategic reciprocity in world politics*. Chicago: The University of Chicago Press.

Gouldner, A. W. (1960). The norm of reciprocity: A preliminary statement. *American Sociological Review, 25*, 161-179.

Greffenius, S. (1990). Measure for measure: Patterns of reciprocity in militarized disputes, 1816-1975. *International Interactions, 16*(2), 137-155.

Heider, F. (1944). Social perception and phenomenal causality. *Psychological Review, 51*, 358-374.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Heiss, J. (1981). Social roles. In M. Rosenberg & R. H. Turner (Eds.), *Social psychology* (pp. 66-93). New York: Basic Books, Inc.

Holsti, O. R., & Rosenau, J. N. (1988). The domestic and foreign policy beliefs of American leaders. *Journal of Conflict Resolution, 32*(2), 248-294.

Howell, L. D. (1979). Comparative study of the WEIS and COPDAB data sets. *International Studies Quarterly, 27*, 149-159.

Hwang, K. K. (1987). Face and favor: The Chinese power game. *American Journal of Sociology, 92*(4), 944-974.

Janis, I. L. (1989). *Crucial decisions: Leadership in policymaking and crisis management*. New York: The Free Press.

Kanouse, D. E. (1972). Language, labeling, and attribution. In Jones, E. E., Kanouse, D. E., Kelley, H. H., Nisbett, R. E., Valins, S. & Weiner, B. (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 121-135). Morristown, NJ: General Learning Press.

Kelley, H. H. (1972a). Attribution in social interaction. In Jones, E. E., Kanouse, D. E., Kelley, H. H., Nisbett, R. E., Valins, S., & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 1-27). Morristown, NJ: General Learning Press.

Kelley, H. H. (1972b). Causal schemata and the attribution process. In Jones, E. E., Kanouse, D. E., Kelley, H. H., Nisbett, R. E., Valins, S., & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 151-174). Morristown, NJ: General Learning Press.

Kenny, D. A., Mohr, C. D., & Levesque, M. J. (2001). A social relations variance partitioning of dyadic behavior. *Psychological Bulletin, 127*(1), 128-141.

Larson, D. W. (1988). The psychology of reciprocity in international relations. *Negotiation Journal, 4*(3), 281-302.

Larson, D. W. (1998). Exchange and reciprocity in international negotiations. *International Negotiation, 3*(2), 121-138.

Laumann, E. O., & Marsden, P. V. (1982). Microstructural analysis in interorganizational systems. *Social networks, 4*, 329-348.

Leng, R. J. (1993). Reciprocating influence strategies in interstate crisis bargaining. *Journal of Conflict Resolution, 37*(1), 3-41.

Leng, R. J. (1998). Reciprocity in recurring crises. *International Negotiation, 3*(2), 197-226.

Lu, L. (1997). Social support, reciprocity, and well being. *The Journal of Social Psychology, 137*(5), 618-628.

Lippitt, R., Polansky, & Rosen, S. (1952). The dynamics of power. *Human Relations, 5*, 37-64.

Litman-Adizes, T. (1977). *An attributional model of depression: Laboratory and clinical investigation.* Unpublished manuscript, University of California, Los Angeles.

Macy, M. W. (1989). Walking out of social traps: A stochastic learning model for the prisoner's dilemma. *Rationality and Society, 1*(2), 197-219.

Martin, J. L., & Ross, H. S. (1996). Do migrating circumstances influence family reactions to physical aggression? *Child Development, 67*, 1455-1466.

Martinko, M. J. (1995). *Attribution theory: An organizational perspective*. Delray Beach, FL: St. Lucie Press.

McClelland, C. A. (1983). Let the user beware. *International Studies Quarterly, 27*, 169-177.

McClelland, C. A. (1978). *World event/interaction survey (WEIS) 1966-1978 (ICPSR 5211), Third Ed.*, Inter-University Consortium for Political and Social Research, Ann Arbor, MI: The University of Michigan.

McDonald, D. M. (1995). Fixing blame in n-person attributions: A social identity model for attributional processes in newly formed cross functional groups. In M. J. Martinko (Ed.), *Attribution theory: An organizational perspective* (pp. 273-288). Delray Beach, FL: St. Lucie Press.

Merritt, R. L., & Zinnes, D. A. (1988). Data development for international research, *International Interactions, 14*(2), 95-100.

Michener, H. A., Vaske, J. J., Schleifer, S. L., Plazewski, J. G., & Chapman, L. J. (1975). Factors affecting concession rate and threat usage in bilateral conflict. *Sociometry, 38*(1), 62-80.

Molm, L. D., Peterson, G., &Takahashi, N. (1999). Power in negotiated and reciprocal exchange. *American Sociological Review, 64*(6), 876-890.

Oetting, E. R. (1999). Primary socialization theory. Developmental stages, spirituality, government institutions, sensation seeking, and theoretical implications. V. *Substance Abuse and Misuse, 34*(7), 947-982.

Patchen, M. (1988). *Resolving disputes between nations: Coercion or conciliation*? Durham: Duke University Press.

Patchen, M. (1991). Conflict and cooperation in US-Soviet relations: What have we learned from the quantitative research? *International Interactions, 17*(2), 127-143.

Patchen, M. (1998). When does reciprocity in the actions of nations occur? *International Negotiation, 3*(2), 171-196.

Patchen, M., & Bogumil, D. D. (1995). Testing alternative models of reciprocity against interaction during the Cold War. *Conflict Management and Peace Science, 14*(2), 163-195.

Patchen, M., & Bogumil, D. D. (1997). Comparative reciprocity during the Cold War. *Peace and Conflict, 3*(1), 37-58.

Parks, C. D., & Komorita, S. S. (1998). Reciprocity research and its implications for the negotiation process. *International Negotiation, 3*(2), 151-169.

Paese, P. W., & Gilin, D. A. (2000). When an adversary is caught telling the truth: Reciprocal cooperation versus self-interest in distributive bargaining. *Personality and Social Psychology Bulletin, 26*(1), 79-90.

Pruitt, D. G. (1969). Stability and sudden change in interpersonal and international affairs. *Journal of Conflict Resolution, 13*(1), 18-38.

Raven, B., Schwarzwald, J., & Koslowsky (1998). Conceptualizing and measuring a power/interaction model of interpersonal influence. *Journal of Applied Social Psychology, 28*(4), 307-332.

Richardson, N. R., Kegley, C. W., & Agnew, A. C. (1981). Symmetry and reciprocity as characteristics of dyadic foreign policy behavior. *Social Science Quarterly, 62*(1), 128-138.

Richman, A. (1975). Issues in the conceptualizations and measurements of events data. *International Interactions, 2*, 19-32.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs, 80*, 1-28.

Shavitt, S., & Fazio, R. H. (1991). Effects of attribute salience on the consistency between attitudes and behavior prediction. *Personality and Social Psychology Bulletin, 17*(5), 507-516.

Sigler, J. H. (1972). Cooperation and conflict in US-USSR-Chinese Relations, 1966-1972: A quantitative analysis. *Peace Research Society Papers, 19*, 107-128.

Singer, E. (1981) Reference groups and social evaluations. In M. Rosenberg & R. H. Turner (Eds.), *Social psychology* (pp. 94-132). New York: Basic Books, Inc.

Smith, C. G. (Ed.) (1971). *Conflict resolution: Contributions of the behavioral sciences*. Notre Dame: University of Notre Dame Press.

Sprecher, S. (1998). Insiders' perspectives on reasons for attraction to a close other. *Social Psychology Quarterly, 61*(4), 287-300.

Starr, H. (1978). "Opportunity" and "Willingness" as ordering concepts in the study of war. *International Interactions, 4*(4), 363-387.

Stern, P. C., & Druckman, D. (2000). *International conflict resolution after the cold war: Committee on international conflict resolution*. National Research Council. Washington, D.C.: National Academy Press.

Stoll, R. J., & McAndrews, W. (1986). Negotiating strategic arms control, 1969-1979. *Journal of Conflict Resolution, 30*(2), 315-326.

Tajfel, H. (1973). Social identity and intergroup behavior. *Social Science Information, 13*, 65-93.

Thibaut, J. (1968). The development of contractual norms in bargaining: Replication and variation. *Journal of Conflict Resolution, 12*(1), 102-112.

Thibaut, J. W., & Kelley, H. H. (1967). *The social psychology of groups*. New York: John Wiley and Sons.

van Tilburg, T., Van Sonderen, E., & Ormel, J. (1991). The measurement of reciprocity in ego-centered networks of personal relationships: A comparison of various indices. *Social Psychology Quarterly, 54*, 54-66.

Van Wyk, K., & Radloff, S. (1993). Symmetry and reciprocity in South Africa's foreign policy. *Journal of Conflict Resolution, 37*(2), 382-396.

Vincent, J. E. (1983). WEIS vs. COPDAB: Correspondence Problems. *International Studies Quarterly, 27*, 161-168.

Ward, M. D. (1981). Seasonality, reactions, expectations, adaptations, and memory in cooperative and conflictful foreign policy behavior: A research note. *International Studies Quarterly, 8*, 229-246.

Ward, M. D., & Rajmaira, S. (1992). Reciprocity and norms in U.S.-Soviet foreign policy. *Journal of Conflict Resolution, 36*(2), 343-368.

Ward, M. D., & Widmaier, U. (1982). The domestic-international conflict nexus: New evidence and old hypotheses. *International Interactions, 9*(1), 75-101.

Weiner, B. (1972). *Theories of motivation.* Chicago: Rand McNally.

Weiner, B. (1979). A theory of motivation for some classroom experiences. *Journal of Educational Psychology, 71*, 3-25.

Weiner, B. (1986). *An attributional theory of motivation and emotion.* Berlin: Springer-Verlag.

Weiner, B. (1992). *Human motivation: Metaphors, theories, and research.* New bury Park: CA: Sage Publications, Inc.

Weiner, B. (1995). Attribution theory in organizational behavior: A relationship of mutual benefit. In M. J. Martinko (Ed.), *Attribution theory: An organizational perspective* (pp. 3-6). Delray Beach, FL: St. Lucie Press.

Weiner, B. (1998). Discovering general laws of social motivation. In J.G. Adair, D. Belanger, & K. L. Dion (Eds.), *Advances in psychological science, Vol. 1*, (pp. 93-109). East Sussex, UK: Psychology Press.

Weiner, B., Frieze, I. H., Kukla, A., Reed, L., Rest, S., & Rosenbaum, R. M. (1971). *Perceiving the causes of success and failure.* Morristown, NJ: General Learning Press.

Weiner, B., & Litman-Adizes, T. (1980). An attributional, expectancy-value analysis of learned helplessness and depression. In J. Garber & M. E. P. Seligman (Eds.), *Human helplessness* (pp. 35-58). New York: Academic Press.

Weiner, B., Nierenberg, R., & Goldstein, M. (1976). Social learning (locus of control) versus attributional (causal stability) interpretations of expectancy of success. *Journal of Personality, 44*, 52-68.

Wosinka, W., Cialdini, R. B., Barrett, D. W., & Reykowsi, J. (2001). *The practice of social influence in multiple cultures.* Mahwah, NJ: Lawrence Erlbaum Associates.

Zaccaro, S. J., Peterson, C., & Walker, S. (1987). Self-serving attributions for individual and group performance. *Social Psychology Quarterly, 50*(3), 257-263.

Zander, A. (1971). *Motives and goals in groups.* New York: Academic Press.

Predicting Performance in 'Go' Situations: A New Use for the Cognitive Failures Questionnaire?

J. Craig Wallace
Georgia Institute of Technology

Steven J. Kass & Claudia Stanny
University of West Florida

This study examined the predictive ability of the Cognitive Failures Questionnaire (CFQ) on one's likelihood of responding incorrectly to target and non-target stimuli on a sustained attention task. The Sustained Attention to Response Task (SART) is a go/nogo task in which the 'go' aspect of the task requires a response by the participant and the 'nogo' aspect of the task requires participants to inhibit a response. A sample of 151 participants completed the CFQ and the SART. Results indicated that one of the CFQ factors, namely Memory, possesses the ability to predict performance for the 'go' aspect of the SART. However, none of the CFQ factors could predict performance for the 'nogo' aspect of the task. Possible applications for the CFQ and SART are discussed, as well as ideas for future research.

Research on sustained attention and vigilance dates back to 1948 when Mackworth first measured sustained attention using a sample of Navy radar operators. Robertson, Manly, Andrade, Baddeley, and Yiend (1997) defined sustained attention as "the ability to self-sustain mindful, conscious processing of stimuli whose repetitive, non-arousing qualities would otherwise lead to habituation and distraction to other stimuli" (p. 747). Mackworth (1948) and a host of other researchers (e.g., Giambra & Quilter, 1989; Kass, Vodanovich, Stanny, & Taylor, 2001; Manly, Robertson, Galloway, & Hawkins, 1999) have examined sustained attention over the years and found several relationships between performance on sustained attention tasks and a variety of constructs. Some of the more common personality traits found to be associated with sustained attention are boredom, worry, and poor attentional levels (Robertson et al., 1997).

Factors such as those listed above have implications for performance on a variety of on-the-job tasks and create a need to predict ability to

---

Author info: Correspondence should be sent to: J. Craig Wallace, School of Psychology, 274 5th Street, Georgia Institute of Technology Atlanta, GA 30332-0170: E-mail: CraigWallace@mediaone.net

North American Journal of Psychology, 2001, Vol. 3, No. 3, 481-490.

© NAJP

sustain attention. Sustained attention is required for a variety of situations including radar/sonar monitoring, industrial inspection, and vehicle operation (Jerison, 1977). These situations commonly involve monotonous work, multi-tasking, and high stress. For example, air traffic controllers monitor a radar screen for hours and are responsible for the safe delivery of the airplane's passengers. Decrements in sustained attention may result in adverse effects (e.g., plane crash). If an individual's ability to sustain attention could be assessed, the potential for adverse effects might be reduced through appropriate employee selection, placement, and training.

All individuals, at one time or another, may forget to do something that they are reasonably familiar with, such as forgetting a recently introduced name or throwing away something that they meant to keep. Absentminded people tend "to be inattentive to ongoing activity, to lose track of current aims and to become distracted from intended thought or action by salient but (currently) irrelevant stimuli" (Manly et al., 1999, p. 661). Absentminded actions, or rather cognitive failures, are cognitively based mistakes or errors on a task that a person should normally be capable of successfully completing (Martin, 1983). Robertson et al. (1997) stated that it is likely that ability to sustain attention plays a crucial role in an individual's tendency to experience cognitive failure.

Reason (1974) was a pioneer in the area of cognitive failure, and much of the interest in the area originated with his work. For example, he studied the navigational errors of pilots who plotted flight plans by the wrong end of the compass needle. Norman (1980) also examined cognitive failures, but referred to them as slips and defined them as "the performance of an action that was not what was intended" (p. 71). Norman (1980) went on to state that there are three major categories of slips: errors in the formation of the intention, faulty activation of schemas, and false triggering. Cognitive failures, or slips, have been shown to relate to such factors as short-term memory capacity, attention and vigilance level, incidental learning, and distributed attention (Broadbent, Cooper, Fitzgerald, & Parkes, 1982; Pollina, Greene, Tunick, & Puckett, 1992).

Recent research (e.g., Houston, 1989; Larson, Alderton, Neideffer, and Underhill, 1997; Smith, Chappelow, & Belyavin, 1995) has continued to examine cognitive failure. Smith et al. (1995) examined focused attention, categoric search, and cognitive failure. They found evidence that supported earlier findings (e.g., Broadbent et al., 1982), namely that cognitive failure is negatively correlated with speed in categoric search (i.e., those participants more likely to experience cognitive failures took longer to identify the target in the search). Houston (1989) found that self-focused attention might be one of the underlying factors responsible for cognitive failure in that the attention paid to one's behavior may reduce the attentional level that is required by other tasks. Larson et al. (1997) found that cognitive failure was positively correlated with on-the-job accidents in a military setting, as well as automobile accidents (Larson & Merritt, 1991).

Robertson, et al. (1997) and Manly, et al. (1999) found that total cognitive failure scores measured via the Cognitive Failures Questionnaire (CFQ) were negatively related with performance on a sustained attention task, specifically the number of errors in response inhibition on the *Sustained Attention to Response Task* (SART) (r = -.27). In other words, participants scoring high on the CFQ failed to inhibit responses on the SART when required to do so. The SART is a go/nogo task that requires a person to respond to all non-targets and withhold a response for the target stimulus. The SART attempts to measure sustained attention over a short period of time (i.e., 4.5 minutes), whereas other measures of sustained attention such as the Mackworth Clock require much more time to administer (e.g., 45 minutes). Errors of commission on a go/nogo task are the laboratory analog of everyday slips of action due to false triggering. While previous research examining this topic (i.e., Manly, et al., 1999; Robertson, et al., 1997) has found a significant negative correlation, the relationship between the factor structure of the CFQ and SART performance has not been examined.

The examples of cognitive failure discussed above can occur for several reasons. A cognitive failure can occur when an action is triggered inappropriately, such as turning a switch on instead of off during emergency procedures. Alternatively, a plan of thought or action may be sidetracked by intervening stimuli, such as forgetting to take the trash out after receiving a telephone call, or targeting the wrong item, such as using flour instead of sugar in one's coffee (Manly, et al., 1999).

The purpose of the present study was to examine whether an instrument designed to assess one's cognitive failure susceptibility could predict actual number of errors on a task requiring sustained attention. Previous research has examined the relationship between total CFQ scores and SART performance (e.g., Robertson, et al., 1997). However, research has not been conducted examining the relationship, or lack thereof, between the newly identified factor structure of the CFQ (see Wallace, Kass, & Stanny, 2001) and SART performance. Additionally, previous research has not utilized the 'go' aspect of the task as a performance measure. To date, only the 'nogo' aspect of the SART has been used as a performance measure. Lastly, previous studies examining this relationship have done so using a clinical sample that was relatively small. The current study used a much larger sample and consisted of young adults. Based on previous findings, it was hypothesized that CFQ subscale scores would be significant predictors of SART 'go' and 'nogo' performance.

METHOD

Participants

A sample of 151 university students participated in the present study. Participants were psychology students at a small university in the South. Participation was voluntary and the participants were anonymous. The sample was comprised of 45 males and 106 females. The average age of the participants was 21.7 years (*SD* = 5.5). The ethnic breakdown of the sample yielded the following percentages: 77.5 % Caucasian, 11.5% African American, 5% Hispanic American, 4% Asian American, and 2% other or not reported.

Materials

The *Cognitive Failures Questionnaire* (Broadbent et al., 1982) is a 25-item self-report inventory that contains questions regarding failures in memory, perception, and motor function over the past 6 months. All questions are worded in the same direction. The response format utilizes a 5-point Likert scale (0=never, 4=always) with possible total scores ranging from 0 to 100. High scores on the CFQ suggest that the person is prone to committing cognitively based mistakes such as throwing away a book of unused matches and retaining the used match or forgetting what one came to the store to purchase (Broadbent et al., 1982). Test-retest reliability was found to be .82 over a 2-month time interval (Vom Hofe, Mainemarre, & Vannier, 1998), and internal consistency was good (alpha = .89) for the current version (Vom Hofe et al., 1998).

Four interpretable factors for the CFQ were identified by Wallace and his colleagues (2001). These were Memory, Distractibility, Blunders, and Names. The Memory factor consists of items that are consistent with memory failures and forgetfulness, while Distractibility pertains to perceptual aspects of attentional tasks (e.g., Do you read something and find you haven't been thinking about it and must read it again?). The Blunders factor relates to physical mishaps and the Names factor contains items that relate to memory for people's names. Subscale scores suggest that a person who scores high on a particular subscale might be more prone to commit a cognitive failure with regard to that particular subscale, than a person who scores lower on that subscale. For example, a person scoring high on the Memory subscale might be more susceptible to forgetting than a person who scores lower on the Memory subscale.

The *Sustained Attention to Response Task* (SART) (Robertson et al., 1997) is a go/nogo task in which 225 single digit numbers, 25 of each digit (1-9), are visually presented via computer over a 4.3-minute time span. Each digit is displayed for 250 milliseconds, after which a masking stimulus is presented for 900 msec. to prevent the target from remaining visible due to the time required for phosphor decay. The onset-to-onset interval between stimuli is 1,150 msec. The mask for the SART is an "X" encompassed in a 29mm circle. The mask and all digits were presented in white against a black background. The SART task was presented on Pentium based 133 MHz PC computers via DOS running Experimental Run Time System (Beringer, 1996). All monitors used were color (SVGA) and the viewing area measured 14 inches (diagonally). No restrictions were placed on the movement of the participants during the task.

A go/nogo task requires response inhibition by the participant. In other words, when a target appears (e.g., a specified letter or number) the participant should not respond. Additionally, when non-targets appear the participant should respond. The task allows for two types of measures: response, measured by the 'go' aspect of the task, and response inhibition, measured by the 'nogo' aspect. Participants scoring high on the 'go' measure of the SART fail to respond to necessary stimuli. Participants scoring high on the 'nogo' aspect of the SART fail to inhibit a response when required to do so on the task.

All participants were required to press a key on a keyboard in response to all digits except the number 3. All 225 digits (5 repetitions of each digit x 9 digits x 5 font sizes) were presented in random order in a single session. The font size for each digit was randomly selected from font sizes of 48, 72, 94, 100, and 120 points. This was done to offset any effects that may have arisen from searching for a peripheral feature of the target digit. Font sizes equate to a height ranging from 12mm to 29mm. The design of the task is consistent with that of Manly, et al. (1999) and Robertson, et al. (1997). Reaction time and false alarms were recorded. All participants were encouraged to give equal value to precision and rapidity while completing the test. Practice sessions of 18 digits, including 2 target digits, preceded all data recording sessions. This was done to familiarize the participants with the task and reduce errors due to task novelty. Robertson, et al. (1997) have shown that the SART is reliable over a 1-week period (*r* = .76).

Procedure

After consenting to participate, participants completed the CFQ and the SART. All participants completed the CFQ first and the SART second. This procedure was undertaken as it was the desire of the researchers to examine the predictive ability of the CFQ and its factors on SART performance. Participants were then debriefed about the study after completing the SART.

RESULTS

The results of the present study differ from previous findings (e.g., Robertson et al., 1997). However, new insight has been gleaned for the predictive ability of the CFQ (see Table 1 for means and standard deviations).

TABLE 1 Means and Standard Deviations of Cognitive Failures and SART Errors

Measure	*M*	*SD*
CFQ Total	44.08	15.41
CFQ Memory	9.17	4.79
CFQ Distractibility	12.82	4.17
CFQ Blunders	13.91	5.94
CFQ Names	6.56	2.61
SART 'go' errors	12.54	5.61
SART 'nogo' errors	6.62	10.45

Note. CFQ = Cognitive Failures Questionnaire; SART = Sustained Attention to Response Task.

For the 'go' aspect of the task no significant relationship between total CFQ scores and SART performance was observed [$F_{(1, 149)} = 1.59$, $p = .19$]. However, further investigation of the predictive ability of the CFQ factors demonstrated a single significant relationship: the CFQ Memory factor was a significant predictor of 'go' performance ($R = .24$, $p < .05$) and accounted for roughly 6% of the variance in SART 'go' performance scores. In other words, this factor was able to predict SART performance for omission errors when a response was required. No other CFQ factors were significant predictors of SART 'go' performance (see Table 2).

For the 'nogo' aspect of the task no significant relationship was identified between total CFQ scores and SART performance [$F_{(1, 149)} = .01$, $p = .89$]. Additionally, no individual factors were significant predictors of SART 'nogo' performance (see Table 2).

DISCUSSION

The results of the present study question the predictive ability of the CFQ for response inhibition performance measures. However, the CFQ Memory factor was related to SART 'go' performance. This suggests that individuals who score high on the CFQ Memory subscale might be less likely to respond to a stimulus that necessitates a response. This finding is quite interesting because it appears that individuals who can adequately inhibit a response when required (measured by the 'nogo' aspect of the SART) may be less prone to respond when a response is warranted (measured by the 'go' aspect of the SART). For example, a police officer may correctly identify a harmless stimulus and not respond, but fail to respond to a threatening stimulus. This could jeopardize the safety of the officer and others if a situation like this was encountered. Future research might examine the importance of a person's decision and his or her willingness to accept a false alarm and its consequences.

TABLE 2 Prediction of CFQ Subscale Scores on SART 'go' and 'nogo' Performance

	'Go' Performance		'Nogo' Performance	
CFQ Subscales	*Beta*	*t*	*Beta*	*t*
Memory	.29	2.26*	.19	.91
Distractibility	-.01	-.06	-.03	-.29
Blunders	-.21	-1.72	-.18	-1.51
Names	.07	.75	.13	1.42

Note. * $p < .05$. CFQ = Cognitive Failures Questionnaire. SART = Sustained to Response Task.

The current study failed to support the hypothesis that CFQ scores would predict 'nogo' performance on the SART. The findings are inconsistent with other researchers' findings (e.g., Robertson et al., 1997, & Manly et al., 1999) that examined the predictive ability of the CFQ on SART performance (i.e., nogo performance). While the current study failed to produce similar findings, there are numerous possibilities that could help explain previous findings. For example, the sample that was used in the study by Robertson, et al. (1997) only contained 75 participants and the current study used a much larger sample ($N = 151$) in which no relationship was identified.

One limitation of the study may be found in the task itself. It is possible that the SART and other measures of attention (e.g., Mackworth Clock) assess different attributes of attention (i.e., attention for vigilance task vs. attention for go/nogo task), and should be used as part of a battery of tests to identify individuals with a propensity for committing cognitive failures. Another limitation may be that subjects responded quickly to all SART stimuli. In other words, the participants traded off accuracy for increased speed. Perhaps directions that emphasize accuracy over speed might have been able to overcome this problem. Or possibly, participants lacked the level of motivation for completing the task appropriately and some form of motivation might have been used (e.g., a monetary reward) in an effort to encourage participants to do their best by placing more importance on the decision to respond.

Future research should examine the CFQ in light of a different sustained attention task (e.g., Mackworth Clock) in order to examine the predictive ability of the CFQ on other types of sustained attention tasks (e.g., vigilance tasks). Additionally, a computer simulation that targets a more specific environment (e.g., production line, police shooting exercise) could be used as a task that more accurately represents an environment where a person will be required to complete a sustained attention task. A more realistic task might influence participants' decision-making by placing more importance on the task and possible outcomes of the task (e.g., a simulated power-plant meltdown). This may allow researchers and practitioners alike to identify the best-suited candidate for the position that the simulation portrays (e.g., air traffic controller, police officer).

Additionally, a more balanced and representative sample of the desired population (e.g., air traffic controllers) might be more generalizable to a particular population rather the university student.

Overall, the CFQ shows limited potential as a tool that could be used as part of a battery of tasks and tests to assess an individual's ability to respond to stimuli. However, before such a battery can be utilized a better understanding of sustained attention, cognitive failures, and other related constructs is necessary. It is the hope of the authors that future research will be able to provide a more thorough understanding of sustained attention and its relationship to cognitive failures.

REFERENCES

Beringer, J. (1996). Experimental Run Time System (v 3.18) [Computer programming language]. Frankfurt: Berisoft Corporation.

Broadbent, D. E., Cooper, P. F., Fitzgerald, P., & Parkes, K. R. (1982). The cognitive failures questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology, 21,* 1-16.

Giambra, L. M., & Quilter, R. E. (1989). Sex differences in sustained attention across the adult life span. *Journal of Applied Psychology, 74,* 91-95.

Houston, D. (1989). The relationship between cognitive failure and self-focused attention. *British Journal of Clinical Psychology, 28,* 85-86.

Jerison, H. J. (1977). Vigilance: Biology, psychology, theory, and practice. In R. R. Mackie (Ed.), *Vigilance: Theory, Operational Performance and Physiological Correlates* (pp. 27-40). New York: Plenum.

Kass, S. J., Vodanovich, S. J., Stanny, C., & Taylor, T. (2001). Watching the clock: Boredom and vigilance performance. *Perceptual and Motor Skills, 92,* 969-976.

Larson, G. E., Alderton, D. L., Neideffer, M., & Underhill, E. (1997). Further evidence on dimensionality and correlates of the cognitive failures questionnaire. *British Journal of Psychology, 88,* 29-38.

Larson, G. E., & Merritt, C. R. (1991). Can accidents be predicted? An empirical test of the cognitive failures questionnaire. *Applied Psychology: An International Review, 40,* 37-45.

Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology, 1,* 6-21.

Manly, T., Robertson, I. H., Galloway, M., & Hawkins, K. (1999). The absent mind: Further investigations of sustained attention to response. *Neuropsychologia, 37,* 661-670.

Martin, M. (1983). Cognitive failure: Everyday and laboratory performance. *Bulletin of Psychonomic Society, 21,* 97-100.

Norman, D. A. (1980). Categorization of action slips. In M. Venturino (Ed.), *Selected readings in human factors* (pp. 71-85). Santa Monica, CA: Human Factors Society.

Pollina, L. K., Greene, A. L., Tunick, R. H., & Puckett, J. M. (1992). Dimensions of everyday memory in young adulthood. *British Journal of Psychology, 83,* 305-321.

Reason, J. T. (1974). *Man in motion: The psychology of travel.* New York: Walker.

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). 'Oops!:' Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia, 35,* 747-758.

Smith, A., Chappelow, J., & Belyavin, A. (1995). Cognitive failures, focused attention and categoric search. *Applied Cognitive Psychology, 9,* S115-S126.

Vom Hofe, A., Mainemarre, G., & Vannier, L. (1998). Sensitivity to everyday failures and cognitive inhibition: Are they related? *European Review of Applied Psychology, 48,* 49-55.

Wallace, J. C., Kass, S. J., & Stanny, C. (2001). *Cognitive failures questionnaire revisited: Correlates and dimensions.* Manuscript under review.

The Development of an Attitudes Towards the Sense of Smell Questionnaire (SoSQ) and a Comparison of Different Professions' Responses

G Neil Martin, Feyishola Apena,
Zinath Chaudry, Zelda Mulligan
Claire Nixon
Middlesex University - England

A number of fictional and scientific accounts emphasise the importance of the sense of smell to humans, but no study has systematically examined people's attitudes and beliefs about the olfactory sense. An *Attitudes Towards the Sense of Smell Questionnaire* was constructed to measure people's beliefs concerning the sense of smell, its importance and its uses. It was administered to 105 participants from occupations in which odour is considered important to varying degrees (17 bakers, 19 florists, 17 chefs, 52 controls). Respondents made self-ratings of their olfactory sensitivity and efficiency, their ability to detect and their ability to identify odour. Principal components analysis produced a four factor solution. These factors were: Liking for People, Places and Objects Associated with Odour, Emotional Responses to Odour, Dispensibility of Odour and the Uses and Efficacy of Odour. Each occupation's response on the sensitivity, efficiency, ability to detect and ability to identify odour measures and for each factor was analysed by a 4 (group) x 2 (sex) x 2 (age) independent groups ANOVA. Older florists rated their sense of smell as being more efficient than did younger florists and chefs rated their ability to identify smells as significantly better than that of florists and controls. No significant main effect of group was found for any factor but a significant sex x group interaction was observed for the Emotional Responses to Odour factor: women in the control group regarded sense of smell as being capable of generating greater emotion than did men. Possible clinical applications are suggested.

Popular journalism and the writings of some of our more celebrated literary artists eulogise the capacity of our sense of smell to evoke profound emotions and thoughts. Kipling, for example, famously remarked that smells were surer than sounds or sights to make your heart strings crack, and a myriad poets and novelists such as Suskind (in *Perfume*), Joyce (in *Portrait of An Artist*), Baudelaire (in *The Perfume*), Sheckley (in the *Alchemical Marriage of Alistair Crompton*) and Proust

---

Author info: Correspondence should be sent to: Dr. Neil Martin, School of Social Science, Middlesex University, LondonEN3 4SF, England. U.K.

North American Journal of Psychology, 2001, Vol. 3, No. 3, 491-502.
© NAJP

(in *A Recherche du Temps Perdu)* have alluded to the ability of smell to evoke powerful emotions and generate memories from the distant past. Newspaper and magazine articles consistently highlight the powers of smell to engage our thinking and to alter our mood (e.g., Hope, 1996; Cripps, 1998).

Yet, the sense of smell seems to occupy a relatively inconsequential place in everyday life when compared to other senses. People are generally very poor at identifying even very familiar odours and they routinely regard the sense of smell as being of lesser importance than other senses (Engen, 1987). It is this view of olfaction that has led to its being viewed as the Cinderella of the Senses. We are aware that it plays an important role in the perception of food flavour (although its role is underestimated), and we enjoy products that are scented -even rating them as being of better quality than unscented items (Spangenberg, Crowley & Henderson, 1996).

Empirical research, however, has established that exposure to ambient or more directly presented odours can have significant, positive and negative effects on human behaviour (Martin, 1999). The presence of a pleasant scent has been found to: increase helping behaviour (Baron, 1997); improve anagram formation (Baron & Thomley, 1994); decrease brain activity associated with attention (Martin, 1998); create a more positive impression of a casually-dressed female job interviewee but create a more negative impression of a well-dressed woman (Baron, 1983); increase or decrease visual vigilance (Gould & Martin, 2001; Warm, Dember, & Parasuraman, 1991); improve performance on difficult tasks (Martin, Barrett & Watson, 2001); impair mental arithmetic performance (Ludvigson & Rottman, 1989); temporarily enhance feelings of good health (Martin, 1996); and enhance recall when the odour is present at learning and retrieval (Herz, 1997). Exposure to unpleasant scents has typically negative consequences, including the impairment of mood and cognitive task performance (Rotton, 1983; Schiffman, Sattely-Miler, Suggs, & Graham, 1995).

Research also suggests that odours are capable of generating more emotional memories than are words or sights (Herz & Cupchik, 1992; 1995) and are effective at prompting memories of events experienced years previously (Aggleton & Waskett, 1999; Chu & Downes, 2000).

These two lines of research - one indicating that exposure to ambient odour can alter behaviour significantly in unpredictable ways, the other indicating that we do not hold the sense of smell in high regard - suggest that our attitudes toward, and our lay knowledge of the use and consequences of smell are areas that researchers may productively explore. The study of people's attitudes towards odour, its uses and its perceived importance may reveal (i) how this sense is generally viewed and (ii) how or if people's views are consistent with the olfactory psychology literature.

No study to date has systematically examined people's beliefs about and attitudes towards the sense of smell, although there have been attempts at surveying people's responses to various aspects of smell. Wrzesniewski, McCauley and Rozin (1999), for example, undertook a validity study of an Affective Impact of Odour scale for liking of places, people and objects. As part of this study 116 American and 336 Flemish Belgian students responded to items such as "When you like new food, is it partly because you like the smell?" They found that people were inclined to respond positively to such questions, indicating that respondents sometimes -rather than often- liked new places, people, cosmetic products, and so on because of their odours. Although a questionnaire-based study, the authors did not report any factor analysis of the questionnaire items.

Wrzesniewski *et al.*'s (1999) results were more modest than those reported by a survey undertaken by the Olfactory Research Fund (ORF, 1997). Its website survey asked visitors a number of questions regarding the importance of smell to people's everyday lives (such as how important the sense of smell was to daily relationships with important others; which emotions and behaviours were capable of being affected by smell). Eighty per cent of their respondents (overall *N*>350) reported using environmental fragrances, 68% of those who used fragrances said they did so to make themselves feel better and 56% responded that they used scents to enhance their well-being. The sense of smell was considered more important for daily relationships with the opposite sex and spouse than with family, friends or co-workers. Love was the emotion regarded as being most affected by the sense of smell (anger was regarded as being the least affected).

The National Geographic Smell Survey undertaken in 1986 (Gibbons, 1986) and reported in 1997 (Gilbert & Wysocki, 1997) also examined people's responses to smell, concluding that the stronger the odour, the more likely it was to generate an emotional memory.

Although the ORF report was a survey website report with no statistical analysis, these data indicate that people hold positive and stereotypical views of the uses of the sense of smell. Both this and the data from Wrzesniewski et al. (1999) studies highlight areas in which people regard the sense of smell as being important. Other areas, however, such as the use of smell in food, retail and health were not examined in any depth and both studies employed different measures and examined different areas of olfaction.

The present study was designed to explore attitudes and beliefs about the sense of smell systematically and comprehensively. We constructed a questionnaire to examine people's responses to statements regarding the importance of smell, the use of smell and the effect of smell on behaviour (from attraction and health, to marketing and thinking). Additionally, the responses of various olfactory-related professions to the measure were also investigated. Individuals with a professional interest in smell such as wine-tasters have been shown to be better identifiers and discriminators of odour than are untrained individuals who do not routinely use their sense of smell in their work (Bende & Nordin, 1997). Expert wine tasters are also likely to use more analytic rather than holistic descriptors of smell than are novices (Chollet & Valentin, 2000). Similarly, chefs have been found to be relatively successful identifiers and discriminators of various smells, although the data for chefs are variable (Hirsch, 1990). The present study, therefore, compared the responses of individuals from professions which regularly rely on the sense of smell -bakers, chefs and florists - on the factors extracted from the SoSQ questionnaire as well as other self-report measures. Two variables known to influence olfactory perception are age and sex (cf. Doty & Kobal, 1995) and these were also included as variables in the present study. It was predicted that a group that routinely uses the sense of smell in their work - specifically, chefs- would express a more positive attitude toward the sense of smell than would participants whose occupations are not related to olfaction or participants whose occupations were olfaction-related, but not as dependent on the sense of smell.

METHOD

Participants

One hundred and ten persons participated in the study. Of this sample, 17 were bakers, 19 were florists, 17 were chefs, three were herbalists, two were wine tasters and 52 belonged to other non-olfactory/gustatory professions. The herbalists and wine tasters were removed from inferential analysis because of small group sizes. Of the remaining 105 participants, 51 were men, 54 were women. The mean age was 29.5 years (*SD*= 5.24).

Procedure and materials

Questionnaire Construction. Based on statements reported by Wrzesniewski et al. (1999), the Olfactory Research Fund's website survey (OFC, 1997) and the ideas and thoughts generated in semi-structured interviews with a pilot group of six members of the general public, items were generated which formed the basis of an Attitudes Towards the Sense of Smell Questionnaire (SoSQ). These thoughts and ideas focused on four themes: the perceived importance of the sense of smell, the role of olfaction in health, its use in everyday life and its ability to alter emotion and thinking. Forty six statements were collated and read for sense and accuracy by four independent raters. Any nonsensical or inaccurately presented statements were revised or corrected.

The questionnaire took the form of statements with which respondents agreed or disagreed on a 7-point scale, where '1' indicated strong disagreement and '7' indicated strong agreement. Examples of such statements include, 'When I like a new place, it is partly because I like its odour', 'I am more likely to spend more time in a shop if it has a pleasant smell', 'Some smells can increase feelings of well-being' and so on.

In addition to these statements, the questionnaire also requested information concerning the respondents age, sex and occupation and asked respondents to rate (on 7-point scales) their sense of smell along various dimensions. They were asked how sensitive their sense of smell was ('1' = not at all sensitive; '7'= very sensitive), how efficient it was ('1' = not at all efficient; '7' = very efficient), how well they were able to detect odour ('1' = not at all effective at detecting odour; '7' = very effective at detecting odour), and how well they were able to identify odour ('1' = cannot identify odours very well; '7' = can identify odours very well). The purpose of these self-reports was to measure how participants regarded their olfactory function (i.e., their beliefs about their olfactory ability) rather than to determine their olfactory function *per se*.

Questionnaire distribution. The questionnaire was administered to groups thought to use their sense of smell routinely in their work (cooking, baking, selling flowers) and for whom the sense of smell may be important. The questionnaire was also administered to a control sample who did not routinely use the sense of smell or taste in their work. Participants were recruited either on campus at a London university or via contact with employers of chefs, bakers and florists in the North London area.

RESULTS

Statistical analysis of the data was undertaken in two stages. The first was an exploratory factor analysis to determine the factor structure of the questionnaire. Given the nature and theme of the items in the questionnaire, a distinct factor structure was predicted and these included factors related to liking for odours, the use of odour and the importance of odour. A reliability analysis was also undertaken to omit unreliable items. The second stage involved an inferential analysis of group differences for each factor extracted.

A principal components VARIMAX factor analysis produced a four factor solution, explaining 45.9% of total variance. Table 1 lists all items and factor loadings for each factor, after item analysis.

TABLE 1 Items Loading on Each of the Four Factors

Item	Factor Loading
I may come to like a perfume solely because it is associated with someone I like	.70
I may come to like an odour solely because it is associated with someone I like	.69
When I like a new place, it is partly because I like the odours	.69
I am more likely to spend time in a shop if it has a pleasant smell	.65
I am more likely to enjoy a meal if I like the way it smells	.65
I may dislike an odour solely because it is associated with someone I dislike	.61
I may dislike a perfume solely because it is associated with someone I dislike	.60
When I like food, it is partly because I like the smell	.55
If I don't like a new/cosmetic/health product, it is partly because of the smell	.54
When I don't like a new food it is partly because I don't like the smell	.54
When I like a new person , it is partly because I like their smell	.52
The smell of food may make me feel hungry even if I have just eaten	.52
If I don't like a person I've just met, it is partly because I don't like their smell	.51
I use differently scented cosmetic products according to the mood I am in	.48
When I don't like a new place, it is partly because I don't like the odours there	.46
Certain smells can evoke pleasant memories	.70
Certain smells can evoke unpleasant memories	.69
Smells can evoke feelings of happiness and joy	.68
Smells can evoke feelings of sadness	.66
Some smells can increase feelings of well-being	.64
Some smells can reduce stress levels	.59
Aromatherapy can be used to induce relaxation or energy	.59
Aromatherapy can be used to treat specific illnesses	.52
Some smells remind me of my childhood	.49
I would rather lose my ability to smell than my ability to hear	..93
I would rather lose my ability to smell than my ability to feel	..91
I would rather lose my ability to smell than my sight in one eye	.89
I would rather lose my ability to smell than my ability to see	.88
I would rather lose my ability to smell than my hearing in one ear	.87
I would rather lose my ability to smell than my ability to taste	.84
My sense of smell is just as important as my other senses	.73
Smell is an important part of my life	.63
Certain smells affect productivity in the work place	.57
Aromatherapy can be used to treat specific illnesses	.52
My ability to appreciate food flavour relies mainly on the sense of smell	.49
Aromatherapy can be used to induce relaxation or energy levels	.48

Factor 1 (eigenvalue= 9.58) accounted for 20.8% of the variance and initially comprised 23 items. After item analysis, reliability was good

BEST COPY AVAILABLE

(Cronbach's alpha = .89). Items with factor loadings lower than .4 were deleted. The reliable items remaining were 20. The items appear to reflect the liking for people, places and objects associated with odour [L (iking) F (or) O (dour)].

The second Factor (eigenvalue = 4.89) accounted for 10.6% of the variance and initially comprised 17 items. Items with factor loadings lower than .4 were deleted and reliability was good (Cronbach's alpha =.87). The nine remaining reliable items seem to reflect emotional responses to odour (ERO).

Factor 3 (eigenvalue = 3.73) accounted for 8.1% of the variance and comprised six items. These were the items which related to the dispensibility of the sense of smell (DSS), that is, the sense's relative importance when compared with other senses. Item analysis found no unreliable items (Cronbach's alpha = .95).

Factor 4 (eigenvalue = 2.94) accounted for 6.4% of the variance and comprises six items (Cronbach's alpha = .81). The items seem to reflect uses and efficacy of odour (UEOO).

Descriptive and inferential statistics

Sensitivity, efficiency, detection and identification measures. An initial analysis of the groups' ability to identify and detect smell, their ratings of their sense of smell's efficiency and sensitivity was undertaken in a 4 (group) x 2 (sex) x 2 (age) independent groups ANOVA. Groups were: baker, chef, florist, control. Each group was divided into those over thirty and those under thirty. Four, three-way ANOVAs were conducted, one for each dependent variable (sensitivity, efficiency, detection, identification).

Table 2 shows the overall mean responses (and standard deviations) of each group. These descriptive statistics seem to indicate that the chefs, more than the other groups, rated their sense of smell more highly and as being more effective and sensitive.

Analysis of the sensitivity and detection data indicted no significant main effect nor any significant interactions. However, analysis of the efficiency rating produced no main effect but an interaction between age and group [$F_{(3, 87)} = 2.72$, $p < .05$]. Tests of simple effects showed that this interaction was attributable to the florists ($p < .05$). Older florists rated their sense of smell as being more efficient than did the younger florists. Analysis of the identification data showed a significant main effect of group [$F_{(3, 89)} = 4.83$, $p < .05$). Bonferroni post-hoc tests indicated that chefs rated their ability to identify smells as being significantly better than that of florists and controls ($p < .02$).

Inferential analysis of factors. The overall mean responses (with standard deviations) of each group for each of the four factors are seen in Table 3.

TABLE 2 Overall Means and SDs for Each Group on Each of Four Measures

	Sensitivity	Efficiency	Detection	Identification
Control	5.04 (1.37)	5.22 (1.08)	5.24 (1.24)	4.69 (1.16)
Chefs	5.82 (1.19)	6.18 (0.95)	6.06 (1.97)	6.12 (1.67)
Florists	5.33 (1.12)	5.58 (1.39)	5.63 (1.12)	5.05 (0.91)
Bakers	5.53 (0.72)	5.59 (0.79)	5.47 (0.94)	5.41 (1.12)
Groups Combined	5.33 (1.25)	5.50 (1.24)	5.49 (1.19)	5.10 (1.22)

A series of 4 (group) x 2 (sex) x 2 (age) independent groups ANOVAs were performed for each factor. There was no significant main effect for any factor but a significant sex x occupation interaction emerged for Factor 2 [$F_{(3,\,89)} = 2.683$, $p < .05$]. Tests of simple effects showed that the interaction was attributable to women in the control group rating their emotional responses to smells as being stronger than those of men ($p < .05$).

TABLE 3 Overall Means and SDs for Each Group on Each of Four Factors

	Liking	Emotion	Dispensibility	Uses of
Control	4.50 (1.01)	5.25 (0.94)	5.25 (1.82)	4.87 (1.14)
Chefs	4.49 (1.25)	5.51 (1.10)	5.16 (2.33)	5.51 (1.00)
Florists	4.76 (0.82)	5.44 (0.69)	4.90 (1.52)	4.94 (0.79)
Bakers	4.58 (0.63)	5.24 (0.91)	5.48 (1.62)	5.16 (0.75)
Groups Combined	4.56 (0.96)	5.32 (0.92)	5.21 (1.81)	5.03 (1.02)

DISCUSSION

The present study was designed to develop and construct a questionnaire that would explore and investigate people's attitudes and beliefs about the sense of smell. A principal components analysis produced a four-factor solution explaining 45% of the overall variance. The items suggested a clear factor structure, and reflect Liking for People, Places and Objects associated with Odour (LFO), Emotional Response to Odour (ERO), Dispensibility of the Sense of Smell (DSS) and the Uses and Efficacy of Odour (UEOO).

It is interesting to note that for three of the factors (LFO, ERO and UEOO), respondents were positively disposed towards the sense of smell.

All of the groups (men, women; bakers, chefs, florists, controls) on average agreed that odours could enhance liking for items/places/people associated with them, that odours were capable of evoking emotion in specific contexts, and that odours can be used to change various behaviours. However, when asked whether they would be more willing to relinquish their sense of smell than any of the other senses, a different pattern emerged, with all groups, on average, indicating that they would rather lose their sense of smell than their ability to see or hear (but not touch). This highlights an interesting feature regarding people's attitudes to smell: it is regarded positively and as being effective in changing behaviour but it is not considered to be more important than the other major senses.

Supporting the validity of a sub-scale published in an earlier study, the items in the LFO scale correspond closely to Wrzesriewski et al.'s (1999) Affective Impact of Odour. In fact, the LFO scale includes all items contained in Wrzesriewski et al.'s (1999) subscale. The rationale for expanding the number of items included in the questionnaire appears to be validated, however, by the inclusion of eight other items that also appear to be measuring liking for things, people and places. The results provide support for a general factor that represents people's liking for odour and items, people and places associated with them.

When group differences were analysed, few differences emerged for any Factor. This was unexpected given the evidence indicating that specific professionals who use odour in their work are better at detecting and identifying odours than are non-users (Bende & Nordin, 1997; Chollet & Valentin, 2000). Perhaps these Factors are not related to such distinctions, given that responses to the subjective rating measures (sensitivity and ability to identify) showed distinct group differences. Although the ability to detect and discriminate between odours typically declines with age (see Doty & Kobal, 1995), older florists rated their sense of smell to be more efficient compared with that of younger florists. Given the absence of a main effect of age, perhaps this finding is typical of the current group only. However, a plausible explanation for this finding is that the older florists, having accrued more experience in their profession, regard this experience as having enhanced their olfactory efficiency. The other group difference concerned chefs' ratings of their ability to identify odours: this was significantly higher than that of florists and controls (but not bakers), possibly because the identification of smell is crucial to their professional success.

While no significant main effect of group was found for any Factor, there was a significant interaction between sex and group. Women in the control group were significantly more likely find odours emotionally provocative than were men. Evidence suggests that women are more sensitive to changes in odour concentration and are better at identifying odours than are men (Doty, Appelbaum, Zusho, & Settle, 1985; Doty, 1991). The sex difference observed for the controls seems to support other, controversial evidence showing that women respond more emotionally to smells than do men and that smells evoke more emotionally-hued memories in women (Herz & Cupchik, 1992).

Future research may use this new measure of attitudes toward the sense of smell to address more specific, research questions such as whether cross-cultural differences, age differences and clinical differences exist in people's attitudes to the olfactory sense. Another, more practical application of the measure may exist in clinical settings. The measure may usefully characterise the responses of those individuals with chemical sensitivity, for example, or those who have completely or partially lost their sense of smell. This is especially important given that the present study and others (e.g., Wrzesriewski, et al., 1999) show that we value other senses more. Information from such measures as the SoSQ could provide a rich source of information to clinicians responsible for individuals who have lost their sense of smell through disease or accident.

Finally, from a methodological standpoint, it would be useful to test the validity and reliability of this and other measures further in these contexts and to compare responses on this measure directly with those from the *Affective Impact of Odour Scale* (Wrzesriewski, et al., 1999).

REFERENCES

Aggleton, J.P. & Waskett, L. (1999). The ability of odours to serve as state-dependent cues for real world memories: Can Viking smells aid the recall of Viking experience? *British Journal of Psychology, 90*, 1-8.

Baron, R.A. (1997). The sweet smell of...helping: Effects of pleasant ambient fragrance on prosocial behaviour in shopping malls. *Personality and Social Psychology Bulletin, 23*, 5, 498-503.

Baron, R.A., & Thomley, J. (1994). A whiff of reality: Positive affect as a potential mediator of the effects of pleasant fragrances on task performance and helping. *Environment and Behaviour, 26*, 766-784.

Bende, M., & Nordin, S. (1997). Perceptual learning in olfaction: Professional wine tasters versus controls. *Physiology & Behaviour, 62*, 5, 1065-1070.

Chollet, S., & Valentin, D. (2000). Expertise level and odour perception: What can we learn from red burgundy wines? *Annee Psychologique, 100*, 1, 11-36.

Chu, S., & Downes, J.J. (2000). Long live Proust: The odour-cued autobiographical memory bump. *Cognition, 75*, 2, B41-B50..

Cripps, E. (1998). Scent of a woman. *Independent On Sunday*, 5 July.

Doty, R.L. (1991). Psychophysical measurement of odor perception in humans. In D.G. Laing, R.L. Doty and W. Breipohl (Eds.). *The human sense of smell.* Berlin: Springer.

Doty, R.L., Appelbaum, S., Zusho, H., & Settle, R.G. (1985). Sex differences in odour identification ability: A cross-cultural analysis. *Neuropsychologia, 23,* 667-672.

Doty, R.L., & Kobal, G. (1995). Current trends in the measurement of olfactory function. In R.L. Doty (Ed.), *Handbook of olfaction and gustation* (pp191-226). New York: Marcel Dekker.

Ehrlichman, H., & Halpern, J.N. (1988). Affect and memory: Effects of pleasant and unpleasant odors on retrieval of happy and unhappy memories. *Journal of Personality and Social Psychology*, 55, 5, 769-779.

Engen, T. (1987). Remembering odours and their names. *American Scientist*, 497-503.

Gibbons, L. (1986). The intimate sense of smell. *National Geographic*, September, 324-360.

Gilbert, A.N., & Wysocki, C.J. (1997). The National Geographic Smell Survey Results. *National Geographic*, October, 514- 525.

Gould, A., & Martin, G.N. (2001). 'A good odour to breathe'? The effect of pleasant ambient odour on human visual vigilance. *Applied Cognitive Psychology, 15*, 2, 225-232.

Herz, R.S. (1997). The effects of cue-distinctiveness on odor-based context dependent memory. *Memory and Cognition, 25,* 375-380.

Herz, R.S., & Cupchik, G.C. (1992). An experimental characterization of odor-evoked memories in humans. *Chemical Senses, 17,* 5, 519-528.

Hirsch, A.R. (1990). Smell and taste: How the culinary experts compare to the rest of us. *Food Technology*, September, 96-101.

Hope, A. (1996). Sweet smells that soften our centres. *The Daily Mail*, 18 December.

Ludvigson, H.W., and Rottman, T.R. (1989). Effects of ambient odors of lavender and cloves on cognition, memory, affect and mood. *Chemical Senses, 14*, 525-536.

Martin, G.N. (1996). Olfactory remediation: Current evidence and possible applications. *Social Science and Medicine, 8*, 63-70.

Martin, G.N. (1998). Human electroencephalographic (EEG) response to olfactory stimulation: Two experiments using the aroma of food. *International Journal of Psychophysiology, 30*, 287-302.

Martin, G.N. (1999). Smell- can we use it to manipulate behaviour? *The RSA Journal, 3l4*, 82-83.

Martin, G.N., Barrett, C., & Watson, M. (2000). The effect of odour pleasantness and task complexity on task performance. Manuscript under revision.

Olfactory Research Fund (1997). Response to website sense of smell survey. http://www.fragrance.org//response.html

Rotton, J. (1983). Affective and cognitive consequences of malodorous pollution. *Basic and Applied Social Psychology, 4*, 171-191.

506

Schiffman, S.S., Sattely-Miller, E.A., Suggs, M.S., & Graham, B.G. (1995). The effect of environmental odors emanating from commercial swine operations on the mood of nearby residents. *Brain Research Bulletin, 37*, 4, 369-375.

Spangenberg, E.R. Crowley, A.E., & Henderson, P.W. (1996). Improving the store environment: Do olfactory cues affect evaluations and behaviours? *Journal of Marketing, 60*, 67-80.

Warm, J.S., Dember, W.N., & Parasuraman, R. (1991). Effects of olfactory stimulation on performance and stress in a visual sustained attention task. *Journal of the Society of Cosmetics and Chemists, 42*, 199-210.

Wrzesniewski, A., McCauley, C., & Rozin, P. (1999). Odor and affect: Individual differences in the impact of odor on liking for places, things and people. *Chemical Senses, 24*, 713-721.

Note: The authors thank Lynn McCutcheon and three reviewers for helpful comments on a previous version of this paper. The authors would also like to thank the participants for their time. Those who would like means, standard deviations, and sample sizes on Tables 2 and 3 for men, women, over 30, and under 30 age groups are invited to correspond with Dr. Martin at the address given at the bottom of the first page of this article.

Testosterone and Male Behavior: Empirical Research with Hamsters Does Not Support the Use of Castration to Deter Human Sexual Aggression

Timothy O. Moore
Clark Atlanta University
and
Neuroscience Institute, Morehouse School of Medicine

The experimental manipulation of gonadal hormones in males has led to a general understanding that androgens (e.g., testosterone) modulate aggressive behavior. In numerous mammalian species, it has been commonly observed that castration reduces aggression and testosterone replacement can restore aggressive behavior. Physiologically, castration can effectively lower endogenous androgens, but the specific effects of castration on aggressive behavior in males is not fully understood. In this article, the role of androgens in aggressive behavior will be examined in both animals and humans. Empirical research from an animal model with hamsters will be reported to determine the efficacy of using castration as a treatment to control human behavior (e.g., sexual aggression). In sum, basic animal research can provide clues to how hormones affect behavior in humans, and the experimental findings presented in this article do not support the utilization of chemical castration to deter deviant sexual aggression in humans.

Animal studies can provide valuable information for behavioral researchers to better understand how sex hormones influence behavior. In males, testosterone is one specific sex hormone that has been implicated in the control of reproductive and aggressive behavior. Testosterone can be classified as a gonadal hormone or androgen, and castration is one experimental technique that can be employed in a laboratory setting to investigate the role of sex hormones on male social behavior. The aim of this article will be to focus on the effects of castration on aggressive behavior in animals as a clue to how efficacious castration is for controlling human behavior. Data from preliminary research with male hamsters will provide the foundation for the views expressed in this article.

Author's Info: Correspondence concerning this article should be sent to Timothy O. Moore, Ph.D., Department of Psychology, Clark Atlanta University, Atlanta, GA 30314

North American Journal of Psychology, 2001, Vol. 3, No. 3, 503-520.

© NAJP

ORGANIZATIONAL AND ACTIVATIONAL EFFECTS OF TESTOSTERONE

Testosterone is an androgen, and it is one of many steroid hormones that has an organizational and activational role in male reproductive physiology (Baum, 1999; Kandel, Schwartz & Jessell, 2000). Other steroid or sex hormones would include progesterone and estrogen, and a simple illustration of the biosynthesis of gonadal hormones in males can be found in Fig.1. Testosterone, progesterone, estradiol and estrogen are found in abundance in the blood of both genders, but during development, there are changes in brain morphology that reflect the presence of different levels of steroid hormones. A small portion of the neurons in the brain have receptors for specific sex steroids, and testosterone receptors are concentrated in genitalia, hair follicles, certain muscle groups, and a few brain regions that regulate male development. As the brain differentiates to form a male or a female, steroid hormones can activate sexually dimorphic nuclei (i.e., parts of the brain that differ in structure and appearance between the two genders) (see Matsumoto, 1999).

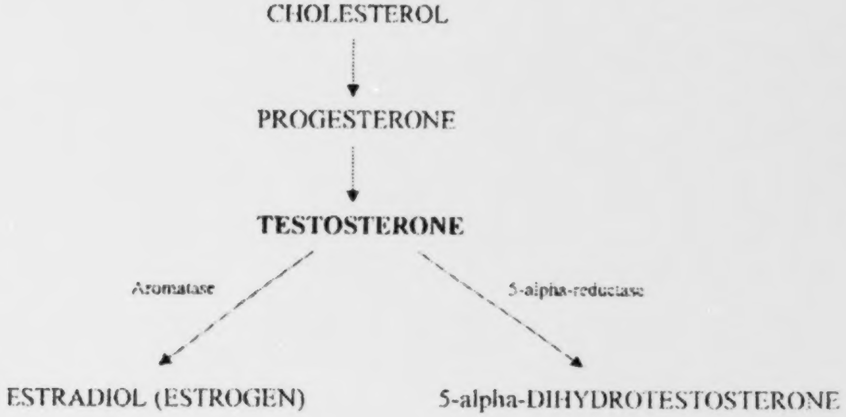


FIGURE 1 Synthesis and metabolism of gonadal hormones in males. Testosterone is synthesized from progesterone, and, in turn, is metabolized in the fetal and adult nervous system into estradiol (via aromatase) or dihydrotestosterone (via 5-alpha-reductase). Each steroid hormone can have important metabolites which are not included in this drawing. An example of a testosterone metabolite involved in aggressive behavior is androstenedione.

Gonadal hormones exert permanent effects on the developing central nervous system and transient effects on the mature adult brain. Whether the effects are permanent or transient, gonadal hormones can act on neurons in at least two ways. One is the classical pathway involving steroid receptors

BEST COPY AVAILABLE

that activate or inhibit specific genes upon binding steroid. This is the mechanism presumed to mediate organizational effects of hormones. Organizational effects of steroids cause the physical changes that separate the genders, such as development of genitalia and musculoskeletal differences, and create central nervous system circuits that program reproductive behaviors. The second way steroids act on neurons involves direct action on the cell membrane via specific receptors. These activational effects are those seen at any time in life to stimulate steroid receptors and temporarily modify sexual, aggressive and other behavioral activities (Kandel et al., 2000).

The effects of gonadal hormones on sexual behavior will be briefly mentioned in this article to present a context for sexual aggression in humans that can lead to deviant social behavior. The motivation and physical expression for sexual and aggressive behavior can be intertwined, and gonadal hormones can influence the expression of both behaviors in males. There is substantial evidence that gonadal hormones affect male reproductive behavior, but the activating effects of gonadal hormones on aggressive behavior in adult mammals deserves attention.

TESTOSTERONE AND MALE BEHAVIOR

In most species, the male is more aggressive than females, and testosterone is the hormonal factor related to this sexually dimorphic behavioral characteristic. In most mammals, the brain in either sex is feminine until it is converted to a masculine form in males through the action of testosterone (Goy & McEwen, 1980). It is well known that during critical stages in neonatal development, testosterone induces organizational effects in the brain that later are expressed in terms of aggressive behavior (Schulkin, 1999). The hormone's early impact means a lifelong tendency toward aggressive behavior (Glickman et al., 1992), and this tendency may be expressed through specific neural circuits. In rodents, for example, there is a neural circuit linked to aggression that includes the medial preoptic and anterior hypothalamic areas, the medial bed nucleus of the stria terminalis, and the medial nucleus of the amygdala (Albert, Jonik & Walsh, 1992; Ferris et al., 1997).

Masculinization in rats can be reversed by removing testosterone from the male during the first 15 days after birth (Schulkin, 1999), so the presence of high levels of testosterone in males has been the sole source of the endocrine connection to aggressive behavior. Other than aggressive behavior, testosterone can influence the central motive state in male rats to search for female rats and then try to mount them under suitable conditions (Beach, 1948). Also, there is evidence that the hormone itself is rewarding to male rats since male rats injected with testosterone returned to the place where the injection occurred more frequently than to others (Alexander, Packard, & Hines, 1994). Moreover, male rats treated with testosterone can discriminate between female rats in estrus and those that are not (Baum, 1992).

Further evidence implicating testosterone and aggressive behavior stems from studies with female animals. In numerous species, it is known that females injected with testosterone during critical stages develop male-like aggressive behavior when they become adults (Schulkin, 1999). They also express male-like characteristics during play-fighting and other sexually dimorphic behaviors. The spotted hyena is one exception to the rule that males are more aggressive than females. Females are both larger and more aggressive than males as a result of being exposed to higher concentrations of testosterone and its metabolites in the womb (Glickman, et al., 1992). Female spotted hyenas have external male-like genitalia and they exhibit rougher play behaviors than do males. The presence of androstenedione, a testosterone metabolite, is higher in female hyenas compared to males, and this hormone plays a key role in the development of aggressive behavior.

CASTRATION AND AGGRESSIVE BEHAVIOR

In 1849, in one of the first experimental investigations with barnyard roosters, Berthold showed that castration can reduce endogenous testosterone and reduce aggressive behavior (see Archer, 1988). Berthold castrated male roosters and then compared their behavior before and after castration. Intact birds were noisy and belligerent and castrates were quiet and non-aggressive. When castrates were given daily injections of testosterone to replace the missing hormone, aggressive behavior reappeared. The consistent finding across numerous species has been that castration in mammals reduces aggression and testosterone replacement can restore aggressive behavior.

CASTRATION AND HAMSTER BEHAVIOR

There are numerous animal models that could be discussed, but extensive preliminary findings will be presented to demonstrate the value of research using hamsters for the topic of gonadal hormones and aggressive behavior. Hamsters are spontaneously aggressive animals, and they are an excellent animal model to study the effects of castration on aggressive behavior (Payne and Swanson, 1970; Swanson, 1967). Although it has been widely accepted over the last three decades that testosterone can modify agonistic behavior in hamsters (Drickamer et al., 1973; Evans & Brain, 1974; Grelk et al., 1974; Payne, 1973, 1974; Payne & Swanson, 1971, 1972; Potegal et al., 1980; Tiefer, 1970; Vandenbergh, 1971; Whitsett, 1975), it is unclear whether gonadal hormones are critical or even necessary

BEST COPY AVAILABLE

for the expression of aggressive behavior in adult male hamsters. Two aggression paradigms that have been used extensively are the resident-intruder model and the neutral cage model. The first paradigm is a familiar setting and the second paradigm is an unfamiliar setting. In both paradigms, hamsters regularly display spontaneous aggressive behavior in the presence of a conspecific. In some studies, castration and the administration of gonadal hormones to castrates can significantly influence the expression of aggression in hamsters using either the resident-intruder model (Payne, 1973, 1974) or the neutral cage paradigm (Grelk et al., 1974; Payne & Swanson, 1971, 1972; Vandenbergh, 1971). In other reports where there was no reduction in male aggressive behavior following castration (Evans & Brain, 1974; Tiefer, 1970; Whitsett, 1975), the neutral cage model was used. In this article, data from both experimental settings will be presented.

In addition to altering aggressive behavior, sexual behavior can be influenced by castration and the removal of circulating androgens. Understanding the reduction in both aggressive and sexual behavior following castration in animals can provide answers to the hormonal mechanisms that may stimulate a human to be sexually aggressive. Other factors such as social context, past history and mental state can profoundly affect a person to exhibit sexual aggression, but it is important to understand the potential influence of male hormones. In studies with hamsters, for example, there is evidence that sexual behavior is reduced after castrating male hamsters and the behavior can be restored with testosterone replacement (Wood & Newman, 1995). Compared to lower animals, sexual behavior in humans is less closely linked to concentrations of gonadal hormones, so it may be more difficult to determine the effects of castration on human behavior. The empirical evidence presented in the next section is focused specifically on male hamster aggression. The potential use of castration to deter repeated sex offenders from committing sexual crimes will be extrapolated from the results.

EXPERIMENTAL FINDINGS FROM THE LABORATORY OF NEUROENDOCRINOLOGY AND BEHAVIOR

Before we can understand the effects of hormones on human behavior, it is necessary that we explore animal research to help us better elucidate the neurohormonal interactions that influence behavior. The following data are from four separate experiments that were conducted in a systematic fashion to determine whether or not gonadal hormones influence aggressive behavior in hamsters. Each experiment was conducted separately over a 24 month period with a different set of hamsters for each experiment, but the data were analyzed collectively in this article to provide a foundation for the ideas discussed here.

Adult male Syrian hamsters (110-125 g) from Charles River Laboratories were isolated for two weeks in our animal colony prior to experimentation. Hamsters were kept on a 14hr:10hr light:dark cycle, and food and water were available *ad lib*. Group-housed hamsters remained group-housed 4-5 per cage to serve as non-aggressive opponents throughout experimentation.

In each of the four experiments, experimental hamsters were prescreened for aggressive behavior with a nonaggressive opponent for five minutes to determine whether or not aggressive behavior would be displayed and to normalize aggressive behavior between experimental animals. If a hamster did not attack a minimum of five times in five minutes, then it was excluded from the experiment. There was a bias to select experimental animals that would display aggressive and not submissive behavior to ensure we could adequately test the effects of castration. Following prescreening, experimental hamsters were castrated or sham castrated. Behavioral testing for each experiment began six weeks post-castration. The outcome of each aggressive encounter produced a dominant (i.e., winner) and a subordinate (i.e., loser) animal. Experimental hamsters were tested in one of the following conditions: (1) resident-intruder model with group-housed opponent; (2) resident-intruder model with isolated opponent; (3) neutral cage model with group-housed opponent; and (4) neutral cage model with isolated opponent. Each experiment was analyzed separately, and all behavioral tests were recorded for five minutes by two independent observers blind to the experimental condition.

To interpret the effects of castration on aggressive behavior, the number of attacks was chosen to represent the display of aggressive behavior. Figure 2 illustrates the number of attacks observed in each social condition. A significant reduction in the number of attacks was only detected in the neutral cage arena with an isolated opponent. An independent t-Test comparing the aggressive behavior of castrates and sham castrates in a neutral cage model with an isolated opponent revealed a statistically significant difference on the number of attacks [$t_{(12)}=2.97$; $p<.01$]. It appears as though castrates were not as successful winning an encounter in a neutral cage with an equally aggressive opponent. On the other hand, there were no significant differences between the number of attacks between castrates and sham castrates in either the resident-intruder model with either type of opponent or the neutral cage model when paired with a group-housed opponent. Overall, castrates fought as frequently as sham castrates in three of the four social conditions. Although different groups of animals were used in each experiment, aggressive behavior was relatively stable for castrates in three of the four conditions, whereas the behavior of sham castrates was more variable across experiments.

FIGURE 2 Number of attacks in five minutes in a resident-intruder model (2A) and a neutral cage model (2B). Castrates (dark bars) and sham castrates (light bars) were paired with either a group-housed opponent or an isolated opponent. A different set of experimental hamsters was used for each condition. In the resident-intruder model, there were no significant differences between groups. In the neutral cage model the number of attacks was significantly lower in castrates vs. sham castrates when paired with an isolated opponent. An asterisk indicates a significant difference between castrates and sham castrates ($p < .01$).

DOES CASTRATION REDUCE MALE AGGRESSION IN ANIMALS?

According to previous research and the findings presented in this article, it is apparent that the effects of castration on male aggression are dependent on the social context of the test situation. Although there is substantive evidence that sex hormones play significant roles in the expression of aggressive behavior in numerous species, other authors reviewing the interaction between hormones and aggression have concluded that this relationship is a highly complex one. The effects of hormonal treatment on aggressive behavior appear to be dependent on interactions with such factors as gender, species, and model of aggressive behavior under consideration (Siegel and Demetrikopoulos, 1993).

Nearly a century after Berthold's (1849) research with roosters, the same effects were observed in other species (Beach, 1948; Guhl, 1961; Young, 1961). In the laboratory, however, the partial retention of androgen-dependent behavioral patterns has received attention (Hart, 1974). Miczek and DeBold (1983) reported that castration pacified only socially naive animals, rats with no prior fighting experience. Rats that were accustomed to defeating intruders, however, remained aggressive even after their testosterone levels were reduced. The findings presented in this article also shed light on the retention of aggressive behavior in experienced aggressive hamsters. The retention is certainly not due to residual amounts of testosterone because the blood levels were non-detectable in castrates versus sham castrates.

There is no simple explanation for why aggressive behavior is retained in castrates or why castrates fought and won in three of the four conditions. Even more puzzling is why castrates lost in the neutral cage paradigm model with an isolated opponent if they could successfully defeat an isolated opponent in a resident-intruder model. Castrates defeated an isolated opponent in a familiar setting, but they were not as successful winning a fight with an isolated opponent in an unfamiliar setting. In contrast to other studies using the neutral cage model (Evans & Brain, 1974; Tiefer, 1970; Whitsett, 1975), a significant reduction in aggression was observed in the neutral cage in this study, but only under specific social conditions.

The castrated hamsters in this study were prescreened for aggressive behavior. It is possible that the experience of behaving aggressively makes the mediating neural substrate less sensitive to androgen changes and the behavior much more resistant to the effects of castration (Hart, 1974). The finding that aggressive behavior is maintained following castration is not an anomalous result. In numerous species, the lack of an effect of castration on reducing aggressive behavior has been observed. For example, Demas and colleagues (1999) demonstrated that castration does not impair aggressive behavior of male prairie voles in either a resident-intruder or a neutral cage model. Moreover, Phoenix, Slob and Goy (1973) observed no evidence of a decline in aggressive responses in adult male rhesus monkeys as much as one year after castration. In rhesus monkeys, neither seasonally nor surgically induced changes in hormone levels altered aggressive behavior (Dixson, 1980; Gordon, Rose and Bernstein, 1976).

One factor that should be considered in studying the effects of castration on aggression is the subject. Barfield, Busch and Wallen (1972) have pointed out that the subject is both a responder and a stimulus to another animal. Rodents primarily use olfactory cues for communication and a significant change in gonadal hormones in castrates could differentially affect the opponent's response to the subject. The effects of androgens on the tendency of the subject to be aggressive is important, but the castrate's capacity to evoke agonistic responses from the other animal is also involved. It is likely that the castrates presented a different stimulus for the isolated opponent in the two different experimental settings. Castrates were able to win encounters in the neutral cage when the subject was group-housed, but they were not as competitive against an isolated opponent in the neutral cage. On the other hand, isolated opponents were expected to be aggressive, and androgens appeared to increase the likelihood of winning an aggressive encounter in a competitive and unfamiliar setting.

In sum, the evidence presented in this article suggests that castration does not always reduce aggressive behavior in animals. We can extrapolate from animal research that the effects of castration in humans may be similarly expressed. Since the same controlled laboratory experiments conducted on animals could not be performed on humans for ethical reasons, experimental research from animal studies can provide insight to how gonadal hormones affect human behavior. Individual differences can exist in animals from experiment to experiment and the factor of individual differences can be even more pronounced in research with humans. Consequently, the effects of gonadal hormones on aggressive behavior are complex and dependent on an interaction between the organism's past experience, the external environment, and endogenous levels of androgens. Next, we will explore the effects of testosterone on aggressive behavior in humans.

TESTOSTERONE AND AGGRESSION IN HUMANS

The nature-nurture debate is a prelude to the assumption that testosterone is responsible for aggressive behavior in humans. On one side of the debate are the biology advocates, who suggest that male hormones are primarily responsible for aggressive behavior. On the other hand, the nurture advocates suggest that males learn aggressiveness and they are the product of a culture that glorifies violence. The evidence for a causal relationship between testosterone and aggression in humans is equivocal. A more prudent position to take is that the social environment in the middle links the endocrine system and the external world to affect an individual's aggressive nature (Niehoff, 1999).

Testosterone levels in boys begin to surge at approximately age ten, rise rapidly over the next several years, and finally begin to plateau at about age fourteen (Susman, Worrall and Murowchick, 1996). During the plateau period, peer group rankings shift daily, with petty quarrels, athletic competitions, academic events, conflicts with parents, and girls offering ample opportunities for boys to be aggressive. Testosterone availability has increased during puberty to handle the challenges associated with these social conflicts. These are logical interpretations of how testosterone may affect human behavior, but there are data which fail to support a relationship between antisocial and/or aggressive behavior and elevated testosterone levels (Nottleman, Susman & Inoff-Germain et al., 1987; Susman, Inoff-Germain & Nottleman, 1987). In another study (Olweus, Mattsson & Low, 1988), boys with higher levels of testosterone were more impatient and irritable, readily frustrated, and quick to take offense. The tendency to start fights, rather than fight back, was less strongly correlated with hormone levels. Hormones may be the most obvious biological event of puberty, but adolescent violence, like adult violence, is more likely to be a product of a brain that has developed a misguided notion of threat than a brain fired up by testosterone (Niehoff, 1999).

The psychoactive effects of testosterone, however, cannot be overlooked. A convincing argument for the effect of testosterone replacement therapy can be found in research with hypogonadal men. Correcting testosterone deficiency in these men results in significant subjective and objective improvement in energy, vitality, and sexual desire and activity (Alexander et al., 1997). Although the effects of replacing physiologic testosterone doses to hypogonadal men is clear, the effects of supraphysiological doses is more complex (Yates, 2000). The illicit use of anabolic-androgenic steroids can provide a source of information about the behavioral effects of increased levels of gonadal hormones in humans.

The rise in the illicit use of testosterone and the synthetic analogs called anabolic-androgenic steroids (AASs) is an example why research is needed on the behavioral effects of gonadal hormones in adult humans. AASs have been used illicitly to improve athletic performance and personal appearance (Buckley et al., 1988; Durant et al., 1993; Yesalis, Kennedy, Kopstein & Bahrke, 1993;). Over the last decade, there have been increasing reports that some AAS users develop marked aggression (Burnett & Kleiman, 1994; Choi, Parrott & Cowan, 1990; Parrott, Choi & Davies, 1994; Pope & Katz, 1988;). Other behavioral changes include mania, hypomania, depressive symptoms and suicidiality during AAS withdrawal. Understanding the effects of testosterone on human behavior is made more complex because individual differences exist and the context in which the testosterone or the AAS is consumed could affect the behavioral response. In a randomized controlled trial with 56 men aged 20 to 50 years of age, a battery of tests were conducted on the effects of supraphysiological doses of testosterone on mood and aggression (Pope, Kouri & Hudson, 2000). The results were variable and not uniform across subjects. Most showed little psychological change, whereas a few developed prominent effects. The authors concluded that the mechanism of these variable reactions remains unclear.

The role of testosterone in male aggression is a public health concern. If testosterone can stimulate people to exhibit higher levels of aggressive behavior, then there may be an escalation of aggressive activity or violent acts committed in society. Therefore, the illicit use of AASs can be problematic. Providing AASs to young people cannot be conducted for ethical reasons, but using animal models with hamsters can provide valuable information

For example, hamsters treated with high doses of AAS during their adolescent development showed heightened measures of aggression in a resident-intruder model (Melloni et al., 1997). Harrison and colleagues (2000) recently examined the hypothesis that exposure to AASs during adolescent development predisposes hamsters to heightened levels of aggressive behavior by influencing specific nuclei in the hypothalamus. Their data suggest that adolescent AAS exposure can increase aggression intensity in hamsters by altering anterior-hypothalamic-arginine vasopressin expression and activity. In conjunction with animal studies, it is necessary to conduct human research to determine the behavioral effects of elevating testosterone with AASs as well as reducing testosterone with chemicals.

IS CHEMICAL CASTRATION A PLAUSIBLE TREATMENT FOR SEX OFFENDERS?

Chemical modification of the sex steroids or their receptors has been developed as a treatment for sex offenders. If sexual aggression is an endocrine problem, then removing the offending hormone must be the solution. It sounds like a logical step to treat sexual aggression, but the animal data presented in this article demonstrates that castration is a complex phenomena.

In a review of hormone treatment of sex offenses (Thibaut, Kuhn, Cordier & Petit, 1998), it has been reported that recent advances in the understanding of factors that are associated with sexual aggression have led to improved methods of treatment. A number of organic treatments which reduce the plasma testosterone levels or decrease androgen effects on target cells is available. These treatments may reduce the chance of re-offending, so the judicial system would invest time and resources into chemical castration as a plausible treatment for sexually aggressive offenders.

Depo-Provera (medroxy-progesterone acetate) was synthesized for use as a contraceptive but was later found to have extensive properties as an androgen antagonist, reducing testosterone levels in males and effecting a state that is the same as reversible chemical castration (Gordon et al., 1970; Meyers et al., 1985). Use of the drug to treat rapists is not supported by psychiatric studies and would appear to be a medical experiment conducted by judges outside of the medical or research establishment (Ayres, 1996). There is controversy over the use of chemical castration in humans and the arguments are based on the myth that testosterone is the hormone for aggression.

Reducing testosterone levels in impulsive or antisocial sex offenders is likely to have an impact only on their sexual behavior because their principal problem is aggression, not sex (Niehoff, 1999). When you take away sex from a rapist who is also a psychopath, you must still treat the psychopathic behavior. Trials with Depo-Provera and other testosterone blocking drugs, involving hundreds of sex offenders, have demonstrated that although these drugs appear to decrease sexual behavior (Zigmond, Finlay and Sved, 1995), they have little or no effect on antisocial aggression (Raine, 1993, 1996).

For some rapists and child molesters, sex is a weapon and the deviant person acts impulsively to assault another person. The aggressive act could be for revenge, power or control. In the case of rape, the deviant behavior is considered to be an act of assault and rape is the gun. Depo-Provera would not effectively treat someone who has the neural circuitry, mentality and social experience for committing aggressive sex acts.

Paraphilias are disorders of sexual deviation and the offenders are motivated by intense, uncontrollable cravings for socially unacceptable sex acts and partners. Sexual behavior is malleable, so treating some paraphilias with Depo-Provera has been successful (Berlin et al., 1991). In more severe cases of paraphilia that involve predatory behavior and even death of the victim(s), Depo-Provera would probably be ineffective. The sexual needs of predatory psychopaths are so inexorably enmeshed with their antisocial aggression that many experts fear that they may be beyond responding to any intervention, chemical or otherwise (Niehoff, 1999).

In 1980, Freund reported on three sex-drive-reducing therapies designed for male patients: 1) pharmacological reduction of circulating testosterone; 2) surgical castration; and 3) deletion of certain brain structures. Since the Freund report, Thibaut et al. (1998) reviewed the current literature and reported that surgical castration and treatment with estrogens or Depo-

Provera are no longer used. The authors advocate a clinical treatment of a gonadotropic releasing hormone analogue (triptoreline) in conjunction with psychotherapy, which seems to benefit offenders who are motivated for treatment.

In sum, the primary effect of castration is to remove endogenous levels of testosterone. As discussed previously, the physiological effects of removing testosterone appear to have varied behavioral effects that probably exist due to the social context of the test situation and individual differences within subjects. In the case of repeated sex offenders, chemical castration is not effective because the problem is primarily aggression and not an abnormal appetite for sex. Chemically castrating humans is not recommended as a primary treatment to control repeated sex offenders such as rapists and child molesters. Compared to chemical castration as a single treatment, incarceration in conjunction with psychotherapy can be used to help a patient manage urges that lead to sexual aggression. Overall, the effects of testosterone are difficult to understand when it is chemically enhanced with AASs as well as surgically or chemically removed with Depo-Provera.

CONCLUSION

It is difficult to predict behavior because it is a challenge to decipher the interplay between internal as well as external factors that affect behavior. Without controlled animal experiments, the specific brain mechanisms involved in producing behavior would be difficult to understand. In this article, we have presented evidence that male hamsters maintain aggressive behavior following castration. They were capable of winning and exhibiting aggressive behavior in specific social conditions. However, when challenged with a potentially aggressive opponent in a neutral arena, castrates were not as successful winning a fight. It appears as though gonadal hormones can affect aggressive behavior in male hamsters but the strength of the effect is dependent on the social context.

The retention of sexual behavior in mammals following castration may be due to the principle that in animals with a more highly evolved forebrain, gonadal hormones are less important for maintaining aggression and sociosexual behaviors (Hart, 1974). In human and nonhuman primates, it is well known that male hormones are an important factor in male sexual behavior. Since human as well as nonhuman primate behavior is controlled by intricate neural circuits and multiple social factors that influence sexual behavior, the chemical control of behavior is less critical as we go up the phylogenetic scale.

Due to the complexity of human life, testosterone's varied effects in humans are probably more difficult to understand versus lower animals.

However, primates, sex and status go hand in hand, and clues to the role of testosterone in male aggression can be found in nonhuman primates. From reviewing nonhuman primate literature, Niehoff (1999) has questioned if testosterone could have more to do with the desire to win versus the desire to kill. In a related finding, we observed that a lack of testosterone in one group of castrated hamsters appeared to decrease their "desire" or motivation to defeat an aggressive conspecific in a neutral cage model. The desire to win is an interesting conceptualization for the role of testosterone because it may relate to competition in the human context. For example, higher levels of testosterone may increase the likelihood of winning a competitive encounter. It does not preclude, however, that a person with high levels of testosterone will always be more aggressive.

In conclusion, our brains are molded by our experiences and some experiences are long-lasting and cannot be modified by a single chemical such as testosterone. Surgical and chemical castration in males can effectively lower gonadal hormones, but a reduction in aggressive behavior following castration is not clearly evident. In both animals and humans, the type of social experience presented before the organism as well as the context of the situation presented before the organism can affect the way gonadal hormones influence behavior.

REFERENCES

Albert, D.J., Jonik, R.H. & Walsh, M.L. (1992). Hormone-dependent aggression in male and female rats: Experiential, hormonal, and neural foundations. *Physiology and Behavior, 16*, 177-192.

Alexander, G.M., Packard, M.G. & Hines, M. (1994). Testosterone has rewarding affective properties in male rats: Implications for the biological basis of sexual motivation. *Behavioral Neuroscience, 108*, 424-428.

Alexander, G.M., Swerdloff, R.S., Wang, C., Davidson, T, McDonald, V., Steiner, B. & Hines, M. (1997). Androgen behavior correlations in hypogonadal men and eugonadal men. *Hormones and Behavior, 31*, 110-119.

Archer, J.A. (1988). *The Behavioural Biology of Aggression*. Cambridge: Cambridge University Press.

Ayres, B. D. (1996). California child molesters face "chemical castration." *The New York Times*, August A1-10.

Barfield, R.J., Busch, D.E. & Wallen, K. (1972). Gonadal influence on agonistic behavior in male domestic rat. *Hormones and Behavior, 3*, 247-259.

Baum, M.J. (1992). Neuroendocrinology of sexual behavior in the male. In J.B. Becker, S.M. Breedlove & and D.Crews (Eds.), *Behavioral Endocrinology*. MIT Press.

Baum, M.J. (1999). Psychosexual development. In M.J. Zigmond, F.E. Bloom, S.C. Lamdis, J.L. Roberts & L.R. Squire (Eds.), *Fundamental Neuroscience* (pp. 1229-1244). San Diego: Academic Press.

Beach, F.A. (1948). *Hormones and Behavior*. New York: Hoeber.

Berlin, F.S., Hunt, W.P., Malin, H.M. et al. (1991). A five-year-plus follow-up survey of criminal recidivism within a treated cohort of 406 pedophiles, 111 exhibitionists, and 109 sexual aggressives: Issues and outcome. *American Journal of Forensic Psychiatry, 12, 5-28.*

Buckley, W.A., Yesalis, C.E., Fried, K.E., Anderson, W., Streit, A. & Wright, J. (1988). Estimated prevalence of anabolic steroid use among male high school seniors. *Journal of the American Medical Association, 260,* 3441-3445.

Burnett, K.F. & Kleinman, M.E. (1994). Psychological characteristics of adolescent steroid users. *Adolescence, 29,* 81-89.

Choi, P.Y.L., Parrott, A.C. & Cowan, D. (1990). High-dose anabolic steroids in strength athletes: Effects upon hostility and aggression. *Human Psychopharmacology,* 5, 349-356.

Demas, G.E., Moffatt, C.A., Drazen, D.L. & Nelson, R.J. (1999). Castration does not inhibit aggressive behavior in adult male prairie voles (*Microtus ochrogaster*). *Physiology and Behavior, 66*(1), 59-62.

Dixson, A.F. (1980). Androgens and aggressive behavior in primates: A review. *Aggressive Behavior, 6,* 37-67.

Drickamer, L.C., Vandenbergh, J.G., & Colby, D.R. (1973). Predictors of dominance in the male golden hamster (*Mesocricetus auratus*). *Animal Behavior, 21,* 557-563.

Durant, R.H., Rickert, V.I., Ashworth, C.S., Newman, C. & Slavens, G. (1993). Use of multiple drugs among adolescents who use anabolic steroids. *New England Journal of Medicine, 328,* 922-926.

Evans, C.M. & Brain, P.F. (1974). Some studies on endocrine influences on aggressive behaviour in the golden hamster (*Mesocricetus auratus* Waterhouse). *Progress in Brain Research, 41,* 473-480.

Ferris, C.F., Melloni, R.H., Koppel, G., Perry, K.W., Fuller, R.W. & Delville, Y. (1997). Vasopressin/serotonin interactions in the anterior hypothalamus control aggressive behavior in golden hamsters. *Journal of Neuroscience, 17,* 4331-40.

Freund, K. (1980). Therapeutic sex drive reduction. *Acta Psychiatric Scandinavian Supplement,* 287(5), 5-38.

Glickman, S.E., Frank, L.G., Pavgi, S. & Licht, P. (1992). Hormonal correlates of masculinization in female spotted hyenas (*Crocuta crocuta*). 1. Infancy to sexual maturity. *Journal of Reproduction,* 95, 451-462.

Gordon, G.G., Southren, A.L., Tochimoto, S., Olivo, J., Altman, K., Rand, J. & Lemberger, L. (1970). Effect of medroxyprogesterone acetate (Provera) on the metabolism and biological activity of testosterone. *Journal of Clinical Endocrinology and Metabolism, 30,* 449-456.

Gordon, T.P., Rose, R.M. & Bernstein, I.S. (1976). Seasonal rhythm in plasma testosterone levels in the rhesus monkey (*Macaca mulatta*): A three year study. *Hormones and Behavior,* 7, 229-243.

Goy, R.W. & McEwen, B.S. (1980). *Sexual Differentiation in the Brain.* MIT Press.

Grelk, D.F., Papson, B.A., Cole, J.E. & Rowe, F.A. (1974). The influence of caging conditions and hormone treatments on fighting in male and female hamsters. *Hormones and Behavior,* 5, 355-366.

Guhl, A.M. (1961). Gonadal hormones and social behavior in infrahuman vertebrates. In W.C. Young (ed.). *Sex and Internal Secretions* (pp. 1240-126), Baltimore: Williams & Wilkins.

Harrison, R.J., Connor, D.F., Nowak, C., Nash, K. & Melloni, R. (2000). Chronic anabolic-androgenic steroid treatment during adolescence increases anterior hypothalamic vasopressin and aggression in intact hamsters. *Psychoneuroendocrinology*, 25, 317-338.

Hart, B.L. (1974). Gonadal androgen and sociosexual behavior of male mammals: A comparative analysis. *Psychological Bulletin, 81*(7), 383-400.

Kandel, E.R., Schwartz, J.H. & Jessell, T.M. (2000). *Principles of Neural Science* (4th ed.). New York: McGraw Hill.

Matsumoto, A. (1999). *Sexual Differentiation of the Brain*. Boca Raton, FL: CRC Press.

Melloni, R.H., Connor, D.F., Hang, P.T.X., Harrison, R.J. & Ferris, C.F. (1997). Anabolic-androgenic steroid exposure during adolescence and aggressive behavior in golden hamsters. *Physiology and Behavior*, 61(3), 359-364.

Meyer, W.J., Walker, P.A., Emory, L.E. & Smith, E.R. (1985). Physical, metabolic, and hormonal effects on men of long-term therapy with medroxyprogesterone acetate. *Fertility and Sterility, 43*, 102-106.

Miczek, K.A. & DeBold, J.F. (1983). Hormone-drug interactions and their influence on aggressive behavior. In B.B. Svare (ed.). *Hormones and Aggressive Behavior*. New York: Plenum Press.

Niehoff, D. (1999). *The Biology of Violence: How Understanding the Brain, Behavior, and Environment can Break the Vicious Circle of Aggression*. New York: The Free Press.

Nottleman, E.D., Susman, E.J., Inoff-Germain, G. et al. (1987). Relationship between adolescent adjustment problems and chronologic age, pubertal stage, and puberty-related serum hormone levels. *Journal of Pediatrics, 110*, 473-480.

Olweus, D., Mattsson, A. & Low, H. (1988). Circulating testosterone levels and aggression in adolescent males: A causal analysis. *Psychosomatic Medicine*, 50, 261-272.

Parrott, A.C., Choi, P.Y.L. & Davies, M. (1994). Anabolic steroid use by amateur athletes: effects upon psychological mood states. *Journal of Sports Medicine and Physical Fitness, 34*, 292-298.

Payne, A.P. & Swanson, H.H. (1970). Agonistic behavior between pairs of hamsters of the same and opposite sex in a neutral observation area. *Behavior, 36*, 259-269.

Payne, A.P.& Swanson, H.H. (1971). The effect of castration and ovarian implantation on aggressive behaviour of male hamsters. *Journal of Endocrinology, 51*, 217-218.

Payne, A.P. & Swanson, H.H. (1972). The effect of sex hormones on the agonistic behavior of the male golden hamster (*Mesocricetus auratus* waterhouse). *Physiology and Behavior,8*, 687-691.

Payne, A.P. (1973). A comparison of the aggressive behaviour of isolated intact and castrated male golden hamsters toward intruders introduced into the home cage. *Physiology and Behavior, 10*, 629-631.

Payne, A.P. (1974). A comparison of the effects of androstenedione, dihydrotestosterone and testosterone propionate on aggression in the castrated male golden hamster. *Physiology and Behavior*, 13, 21-26.

Phoenix, C.H., Slob, A.K. and Goy, R.W. (1973). Effects of castration and replacement therapy on the sexual behavior of adult male rhesuses. *Journal of Comparative and Physiological Psychology*, 84, 472-481.

Pope, H.G. & Katz, D.L. (1988). Affective and psychotic symptoms associated with anabolic steroid use. *American Journal of Psychiatry*, 145, 487-490.

Pope, H.G., Kouri, E.M. & Hudson, J.I. (2000). Effects of supraphysiologic doses of testosterone on mood and aggression in normal men. *Archives in General Psychiatry*, 57, 133-140.

Potegal, M., Blau, A.D., Black, M. & Glusman, M. (1980). Effects of castration of male golden hamsters on their aggression toward a restrained target. *Behavioral Neural Biology*, 29(3), 315-330.

Raine, A. (1993). *The Psychopathology of Crime*. New York: Academic Press.

Raine, A. (1996). Autonomic nervous system factors underlying disinhibited, antisocial, and violent behavior. *Annals of the New York Academy of Sciences*, 794, 46-49.

Schulkin, J. (1999). *The Neuroendocrine Regulation of Behavior*. Cambridge University Press.

Siegel, A. & Demetrikopoulos, M. (1993). Hormones and aggression. In J. Schulkin (ed.). *Hormonally Induced Changes in Mind and Brain* (pp. 99-127). San Diego: Academic Press.

Susman, E.J., Inoff-Germain, G. & Nottleman, E.D. (1987). Hormones, emotional dispositions, and aggressive attributes in young adolescents. *Child Development*, 58, 1114-1134.

Susman, E.J., Worrall, B.K. & Murowchick, E. (1996). Experience and neuroendocrine parameters of development: Aggressive behavior and competencies. In D.M. Stoff & R.B. Cairns (ed.). *Aggression and Violence* (pp. 267-289). Hillsdale, NJ: Erlbaum.

Swanson, H.H. (1967). Effects of pre- and post-pubertal gonadectomy on sex difference in growth, adrenal and pituitary weights in hamsters. *Journal of Endocrinology*, 39, 555-564.

Thibaut, F., Kuhn, J.M., Cordier, B. & Petit, M. (1998). Hormone treatment of sex offenses. *Encephale*, 24(2), 132-137.

Tiefer, L. (1970). Gonadal hormones and mating behavior in the adult golden hamster. *Hormones and Behavior*, 1, 189-202.

Vandenbergh, J.G. (1971). The effects of gonadal hormones on the aggressive behavior of adult golden hamsters (*Mesocricetus auratus*). *Animal Behavior*, 19, 589-594.

Whitsett, J.M. (1975). The development of aggressive and marking behavior in intact and castrated male hamsters. *Hormones and Behavior*, 6, 47-57.

Wood, R.I. & Newman, S.W. (1995). Hormonal influence on neurons of the mating behavior pathway in male hamsters. In P.E. Micevych & R.P. Hammer (Eds.). *Neurobiological Effects of Sex Steroid Hormones* (pp. 3-39). New York: Cambridge University Press.

Yates, W.R. (2000). Testosterone in psychiatry. *Archives in General Psychiatry*, 57, 155-156.

Yesalis, C.E., Kennedy, N.J., Kopstein, A.N. & Bahrke, M.S. (1993). Anabolic-androgenic steroid use in the United States. *Journal of the American Medical Association*, 270, 1217-1221.

Young, W.C. (1961). The hormones and mating behavior. In W.C. Young (ed.). *Sex and Internal Secretions*. Baltimore: Williams & Wilkins.

Zigmond, M.J., Finlay, J.M. & Sved, A.F. (1995). Neurochemical studies of central noradrenergic responses to acute and chronic stress. In M.J. Friedman, D.S. Charney & A.Y. Deutch. *Neurobiological and Clinical Consequences of Stress* (pp. 45-60). Philadelphia: Lippincott-Raven.

Author's Note: This work was supported by NIH (Grant# NS37232) and the Center for Behavioral Neuroscience (NSF-STC IBN-9876754). The author would like to thank Elliott Albers and his personnel from Georgia State University in the Laboratory of Neuroendocrinology and Behavior for assistance with animal research.

Chandra Levy, The Missing Intern: Social, Forensic, and Evolutionary Psychology Insights

Russell Eisenman
University of Texas-Pan American

The case of Chandra Levy, the missing intern, is discussed, with insights drawn from social, forensic, and evolutionary psychology. The media are shown to handle the case in a particular way, emphasizing it, in part, because of the attractiveness--in the broadest sense of the term--of Ms. Levy. The concept of deviance explains a lot about how the case has been presented, with initial views of Congressman Condit being mostly favorable, until he was forced to reveal facts. Ultimately this resulted in his acquiring stigma. The police then handled the case differently. Some ideas are almost never mentioned by the media, such as the possibility of her internship with the Bureau of Prisons putting her in contact with a convict who may have contacted her later, or had a friend contact her. From the standpoint of evolutionary psychology theory, there is nothing unusual in the older congressman being attracted to a young woman, and vice versa.

As this is written, Chandra Levy is still missing. No body has been found, nor has there been any evidence to show that she has chosen, on her own, to disappear. Provided here are some insights into the case drawn from social, forensic, and evolutionary psychology. The way the case is handled in the media reveals important issues confronting the media and our perceptions of public figures.

Sick of the Story, But Fascinated

In many ways, one can become sick of the Chandra Levy story, at least the way it is presented in the media. The main reason is that, early on, the new stories offered so little new information. At times, one could feel like yelling into the television set: "If you have nothing new to say, shut up!" Television reports about the case appeared and reappeared so often, with so little information most of the time. But, it is, per se, an interesting story. It is also interesting in terms of how the media report it, and how some of the people involved in the story deal with it, especially Congressman Gary Condit, Levy's friend and sponsor of her internship with the Federal Bureau of Prisons.

Author info: Correspondence should be sent to: Russell Eisenman, Ph.D. Department of Psychology, University of Texas-Pan American.Edinburg, TX 78539-2999 E-mail: eisenman@panam.edu

North American Journal of Psychology, 2001, Vol. 3, No. 3, 521-528.

© NAJP

The Police and Congressman Condit

The police investigation is proceeding in, apparently, a fairly routine manner, or at least that is what the police want you to think. She is still classified as a missing person, even though there is information to make one think she has been killed, or, perhaps, committed suicide. However, the police do not want to say things that will come back to haunt them--either in court or reputation wise--, so silence is the best tactic for them. They say as little as possible, and pursue the investigation on their own terms. Police often have ways of proceeding that are standard for them, but are virtually unknown to the general public (Arrigo, 2001; Dantzker, 1997, 1998, 1999; McArdle & Erzen, 2001).

Often, the police will do more when a famous person is involved, or a person is involved who has contacts with famous people, or is wealthy, or has wealthy parents, etc. We do not know what is really going on with the police, behind the scenes. We do know, however, that initially, the police publicly took the position that Congressman Condit was not a suspect. He seemed to get special treatment, due to his status. Imagine what would occur if some working class or unemployed person was in his position: likely a former boyfriend, but one who denies it, and denies having much contact with the disappeared person, even though there is reason to think he may be less than totally forthcoming. Would the police say things on television to indicate that they are investigating a missing person case and that even if there was sexual contact that is not what the case is about? That almost never occurs. So, the congressman appears to have received special treatment from the police, at least initially.

The police only started treating him differently when it turned out that despite his early denials, he had had a sexual affair with Chandra Levy and his early reports about limited contact with her turned out not to be true. Then, the police did what they more typically do to people in his position (not his position as Congressman, but his position as the boyfriend of a missing or murdered person). They started treating him more harshly, demanding that he let them search his apartment, for example, and demanding that he take a police or FBI-administered polygraph. What changed? It was, apparently, Condit's admissions that his earlier statements about not being sexually involved with Levy and not having much contact with her were false. Either by omission of certain details of their relationship, or perhaps by issuing misleading statements, Condit apparently sought to minimize his relationship with the young intern. Congressman Condit's admissions appear to have been coerced by Chandra Levy's family, who knew something about the real relationship and, reportedly, threatened to go public if the congressman, himself, did not tell the truth about the relationship, according to television news reports (ABC, CNN Headline News, and others).

Once he admitted his previously denied, or at least not fully admitted, sexual relationship to the police, their image of him changed, and he now was seen as deviant, in the sense of having stigma associated with him and thus being labeled as someone who is not so nice or not so in need of protection. This is how the social perception of deviance often operates (Eisenman, 1991; Goffman, 1963). Something occurs that produces stigma in a person, and then that person is looked at as different, i. e., as deviant, not the same as a normal person (Eisenman, 1991; Goffman, 1963; Heatherton, Kleck, Hebl, & Hull, 2000). Often what occurs to produce stigma is some kind of social labeling. Thus, the person seen as deviant may not be truly guilty of anything really bad, but that person is now perceived that way, due to having been labeled (Eisenman, 1999b).

In fact, even when there is something in the person that makes them truly different, it is still necessary for a social labeling to occur in order for them to be perceived as deviant. Thus, it could be said that social labeling is crucial to the process for deviance to occur. I do not mean, however, to endorse a sociological theory, called social labeling, proponents of which tried to deny that there was anything truly deviant in some people called deviant (Becker, 1963; Lemert, 1972; Schur, 1971). In some cases there is only social labeling and a normal person is seen as deviant, but in other cases, e.g., serious mental illness or serious crime, there is something about the person that is different than most others. In the case of serious mental illness, it may be something biological. In the case of serious criminals, it may be merely their behavior that is different, but that is, nonetheless, a major difference. Most people are not serious, habitual, criminal offenders. Thus deviance can be perceived either because there is something about the person that really *is* different, or because of social labeling of a person who is not different, but is labeled and *perceived* as different (Bryant, 1990; Eisenman, 1991; Goffman, 1961, 1963; Heatherton, Kleck, Hebl, & Hull, 2000).

The police do not want to rule out the possibility that she has run away, perhaps for some of the reasons mentioned here: lover's argument, involvement with a prisoner, unhappiness at not getting her internship extended or not getting a job she wanted, or any other possibility that causes people to run away. And, it may be wise for the police to pretend to trust Congressman Condit even if they do not. By seeming to trust someone the police may elicit a greater degree of cooperation, including having the person talk to them. If the police identify someone as a suspect, that person's attorney would likely tell the person not to speak to the police, or at least to limit what they say.

The Congressman

It would not be unusual, statistically, to think that Congressman Gary Condit might have killed Levy in a lover's dispute. According to the Violence Policy Center (as mentioned on Court TV on August 27, 2001), 57% of female murder victims are killed by boyfriends or husbands. Perhaps she discovered that he was not serious about her, when she hoped he would be. This is a reasonable inference, based on statements made on televised reports by Chandra Levy's aunt, in whom Ms. Levy confided, and statements by Congressman Condit minimizing the relationship. The aunt said that Chandra told her the Congressman had a five-year plan to leave his wife and marry her. Disputes over level of commitment are typical in male-female conflict that occurs from time to time. According to evolutionary psychology, men mainly desire numerous partners who have youth, health, and beauty, in order to spread their genes as much as possible, to insure their genes going into future generations, while women—since they get pregnant—cannot use the same gene-spreading strategy. Instead, women seek men with status and money, who will care for them and their children (Archer, 1996; Buss, 1994, 1999; Dawkins, 1990; Eisenman, 2001; Gaulin & McBurney, 2001).

Buss (1989) tested evolutionary psychology hypotheses in 37 different cultures and found strong support for male-female differences in attraction and other aspects of evolutionary psychology theory, including differences in what makes males vs. females jealous. Buss (1989) found that males are frequently made jealous by any sexual intercourse their partner has, since it could result in her pregnancy and thus his raising a child who will not send his genes into future generations. On the other hand, females are less likely to be jealous over their male partner's sexual intercourse with another woman, but are more likely to become jealous if there is a threat of her man showing commitment to the other woman. Thus, intercourse *per se* makes men jealous, while commitment is the more important issue for female jealousy. From this, it would follow that men would be more often jealous of the partner's having a sexual affair, since any unprotected intercourse is a threat to having his genes go into future generations.

We do not need to hold it against Congressman Condit if he and Ms. Levy were lovers. This would be a very human characteristic, and, from an evolutionary psychology standpoint, would reflect his desire (albeit unconscious; see Eisenman, 2001) to spread his genes as much as possible into future generations. He would not want to admit the sexual affair, as it can hurt his reputation and political career, so he is in a difficult position. But, the police are likely to be concerned that a boyfriend killed his girlfriend--as they routinely think this way in such circumstances--and not let protection of a powerful politician weaken their pursuit of this angle. Since police are, by the nature of their work, a fairly secretive organization, we cannot really know what they think about the case.

Ironically, during the Clinton scandal, also involving an intern, Monica Lewinsky, Congressman Condit called for full disclosure. He said the Starr report on the President should be released to the public, as opposed to coming out a little at a time. He was opposed to having it be "drip, drip, drip." But, now the shoe is on the other foot—his foot--and he is behaving differently. There have been reports that the police have tried to speak with him on a few occasions, and he has attempted to avoid them. He apparently only spoke with them after Chandra Levy's parents met with him, and agreed to stop saying things to the media about him being her lover, if he would meet with the police (although it should be noted that Congressman Condit disputes this, and claims he was cooperative from the beginning). Even in his meetings with the police, he did not reveal until the third of four police interviews that he had an affair with Chandra Levy, according to television news reports.

Why All the Publicity?

However, we should keep in mind that many people disappear. So, why does this story get so much publicity? In part, it is the story of a politician and an intern, bringing to mind the President Clinton-Monica Lewinsky scandal. Also, it is because she is an attractive, young, white woman. If any of those demographic categories about her were different, we would probably not hear so much about it. These four categories—young, attractive, white, female-- are ones that our society values in the case of a victim--so it makes for good television.

The story has some interesting features, but so would the stories of other missing people. However, the media do not necessarily report the truth, but give a slant that appeals to people, such as "real-life" police shows on television, which often present events in ways that differ from what usually occurs (Eisenman, 1999a). For example, they often show a stolen car recovered almost immediately, after an exciting chase. Do you really think, if your car is stolen, that you are likely to get it back soon, and in good shape, and the criminals will be apprehended? But, people enjoy seeing a good ending, so there is profit to be made in presenting things in a positive way, even when it is unrealistic.

A Stranger Could Have Done It

Besides thinking that the congressman was her lover and killed her in a lover's dispute, it is quite possible that a stranger did it. Washington, DC is a very dangerous place to live. The fact that she walked home from her gym workout made her a likely victim to some violent criminal.

The Bureau of Prisons Angle

Chandra Levy worked as an intern for the Bureau of Prisons. Did she have contact with prisoners? This has apparently never been disclosed. Prisoners are often dangerous people. Perhaps one got out and decided to contact her. Or, he passed her name along to a friend. This angle has not been adequately pursued, at least by the media.

Yet, prisoners will sometimes try to contact people they meet in the prison. Also, based on unpublished case studies that have come to my attention, sometimes women fall in love with prisoners, and initiate the contact and future relationships with them, often with disastrous results. These contacts may occur because the prisoner is a master in manipulation, and knows how to make the woman feel good when she is communicating with him. The woman may have a rescue fantasy, in which she imagines or believes that if she is good to the prisoner this will help reform him. Also, since he is in a captive situation, she does not get to see him in a setting where he has the opportunity to misbehave. Once out, any relationship runs the risk of the criminal's real personality coming to the fore, with all the negative things that can mean. Males sometimes get involved with prisoners, too, but the goal is usually strictly sexual, whereas the female involvement with prisoners seems more often predicated on love interests.

Why We Care

What the media have done is make us look at Chandra Levy as a real person, even though, actually, we have only limited insights into her and her life. But, it has shown her to us in a way that we "feel" for her. We are made to feel like we know her. We even see home videos of her and her family. We see her smiling, eating, interacting with other family members. This gives us the false sense that we know her, since these are basic human behaviors that we have all engaged in.

We see her as a fellow human, probably hurt in some unknown way. It arouses empathy. This is done, in large part, because she is young, attractive, and white (possibly, too, because she is female; males are supposed to be tough and take care of themselves, while females are sometimes seen as vulnerable and in need of protection). Youth and attractiveness appeals to almost everyone, while whiteness appeals to the majority of the United States population, which is predominantly white. However, for minorities, they would likely have more interest in a story about someone from their racial or ethnic group.

Conclusion

There are several reasons why the Chandra Levy case is so interesting to many of us. Indeed, the media would not focus on it so much unless they thought it would have a great deal of public appeal. The story of a politician and an intern is interesting for its own sake, and also reminds us of the Clinton-Lewinsky scandal. The demographic characteristics of Ms. Levy-- young, attractive, white, female-- make her attractive to us (in the most general sense of the word "attractive"). The home videos of her, repeatedly shown on television, further humanize her and make us feel as though we know her, even though we do not. The concept of deviance is useful for understanding the changed perceptions of Congressman Condit. Various forensic, social, and evolutionary psychology insights are provided above. To take only one example, from an evolutionary psychology standpoint, it is not unusual that Congressman Condit was attracted to a young intern, and vice versa. Evolutionary psychology theory "sees" this as very normal human behavior, regardless of whether it is socially acceptable.

REFERENCES

Archer, J. (1996). Sex differences in social behavior: Are the social role and evolutionary explanations compatible? *American Psychologist, 51*, 909-917.

Arrigo, B. A. (2001). *Introduction to forensic psychology: Issues and controversies in crime and justice*. San Diego: Academic Press.

Becker, H. (1963). *Outsiders*. New York: Free Press.

Bryant, C. D. (Ed.) (1990). *Deviant behavior: Readings in the sociology of norm violation*. New York: Hemisphere.

Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences, 12*, 1-49.

Buss, D. M. (1994). *The evolution of desire: Strategies of human mating*. New York: Basic Books.

Buss, D. M. (1999). *Evolutionary psychology: The new science of the mind*. Boston: Allyn & Bacon.

Dantzker, M. L. (1997). *Contemporary policing: Personnel, issues and trends*. Woburn, MA: Butterworth-Heinemann.

Dantzker, M. L. (1998). *Police organization and management: Yesterday, today, and tomorrow*. Woburn, MA: Butterworth-Heinemann.

Dantzker, M. L. (Ed.) (1999). *Readings for research methods in criminology and criminal justice*. Woburn, MA: Butterworth-Heinemann.

Dawkins, R. (1990). *The selfish gene*. New York: Oxford University Press.

Eisenman, R. (1991). *From crime to creativity: Psychological and social factors in deviance*. Dubuque, IA: Kendall/Hunt.

Eisenman, R. (1999a). Real-life cop shows on television: Are they really real? *Journal of Information Ethics, 8* (1), 5-9.

Eisenman, R. (1999b). Sexual harassment charges against university faculty: Three case histories. *Journal of Information Ethics, 8* (2), 59-75.

Eisenman, R. (2001). Human sexuality: Answers from evolutionary psychology. *Journal of Evolutionary Psychology*, 22, 53-55.

Gaulin, S. J. C., & McBurney, D. H. (2001). *Psychology: An evolutionary approach*. Upper Saddle River, NJ: Prentice-Hall.

Goffman, E. (1961). *Asylums*. Garden City, NY: Doubleday Anchor.

Goffman, E. (1963). *Stigma: Notes on the management of spoiled identity*. Englewood Cliffs, NJ: Prentice-Hall.

Heatherton, T. F., Kleck, R. E., Hebl, M. R., & Hull, J. G. (Eds.) (2000). *The social psychology of stigma*. New York: Guilford.

Lemert, E. N. (1972). *Human deviance, social problems, and social control* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

McArdle, A., & Erzen, T. (2001). *Zero tolerance: Quality of life, and the new police brutality in New York City*. New York: New York University Press.

Schur, E. (1971). *Labeling deviant behavior*. New York: Harper & Row.

Self-Definition in Psychotherapy: Is it Time to Revisit Self-Perception Theory?

Rostyslaw W. Robak
Pace University

People in psychotherapy define themselves through the narratives and stories they construct about themselves. This post-modern idea is the foundation of much current psychotherapy and family therapy. This notion can be explained by Bem's (1972) self-perception theory. Individuals in therapy observe their own behaviors and define themselves on the basis of their self-perceived behaviors. A six-step process of self-definition based on self-perception in psychotherapy is proposed. When several traditional therapy approaches are examined in this context, self-definition appears to be a key process in psychotherapy. Hypotheses for future research are outlined.

A central process in counseling and psychotherapy is the narration of one's story. Several current approaches to therapy are founded on this premise. The constructivist approach, for example, emphasizes clients' creating their own realities, constructing their life stories, and finding new meanings or themes in their lives (e.g., Franklin & Nurius, 1996; Granvold, 1996; Neimeyer, 1995). In contrast to older "rationalist" cognitive therapies that are based on first-order assumptions about change (i.e., addressing "problem" behaviors directly), constructivist therapies are guided by second-order principles and processes of every individual's constructing his or her preferred realities (Lyddon, 1990). Narrative therapy, a related family therapy approach, emphasizes the construction of story structures in order to give meaning to experiences and to help people see themselves as separate from their problems (e.g., Etchison & Kleist, 2000; Feixas, 1990; Freedman & Combs, 1996; Schwartz, 1999).

Such approaches focus on an underlying process of self-definition. More traditional models of therapeutic change, such as psychoanalytic self psychology (Linden, 1994) and person-centered therapy (Rogers, 1951) also rely on this foundational process. Therapies that emphasize personal narrative necessarily have self-definition as a key process. In therapy, as elsewhere, the definitions of our selves and our self-concepts accrue from the information that we feel defines us. Our self-concepts (perceptions of our own personal characteristics) do not necessarily accurately describe our characteristics, but they do serve to summarize

---

Author info: Correspondence should be sent to: Dr. Rostyslaw W. Robak. Psychology Dept., Pace University, 861 Bedford Rd., Pleasantville, NY 10570.

North American Journal of Psychology, 2001, Vol. 3, No. 3, 529-534.
© NAJP

and organize our assumptions about our personal qualities (McGuire & McGuire, 1988; Rosenberg, 1986).

Self-concepts do change. Persons can have many self-concepts over time (Baumeister, 1998). For instance, people change their self-beliefs based on the self-presentations they make to others, and such changes affect their behaviors and their self-concepts (Fazio, Effrein, & Falender, 1981; Schlenker, 1994). Kelly (2000) has proposed that self-presentation and impression management strategies by psychotherapy clients are crucial to successful psychotherapy. She proposed a four-step model of change in psychotherapy based on a self-presentational process that involves: (1) clients' performing various self-presentations (favorable self-descriptions), (2) therapists' giving feedback based on the clients' self-presentations, (3) such feedback leading to clients' shifts in self-beliefs, and (4) changes in self-beliefs leading to changes in clients' self-concepts. This perspective is significant because it puts self-description at the center of psychotherapy. However, clients' self-presentations (impression management) to an expert audience (the therapist) and therapists' feedback based on impressions from a client-selected presentation may not be enough to capture a key process in psychotherapy.

About thirty years ago Daryl Bem (1972) articulated self-perception theory. According to self-perception theory, people come to know themselves in the same way that they come to know others. They observe their own behaviors in a variety of situations and then they make attributions about their behaviors.

Of course, people are free to make so-called fundamental attribution errors (Ross, 1977). Most of the time we prefer to attribute the causes of behaviors in others to personality traits or internal factors, rather than situational ones. We tend to think that another person's actions are caused by something within, rather than being caused by circumstances, or external, situational factors. The personality traits to which we attribute our own behaviors can be seen as self-definitions. This is especially true in the context of narratives about ourselves in psychotherapy.

It may be useful to outline specifically how self-definition operates in counseling and psychotherapy. Such a model can serve to identify the stages of the process. It should also point out a basis for the reciprocal nature of therapy.

Here is a proposed six-step outline of the self-definition process in counseling and psychotherapy. In a psychotherapy experience, people learn about themselves by observing their own behaviors:

1. They observe and attend to the things they say about themselves in counseling sessions.
2. The therapist encourages new behaviors, including new

recognitions of feelings, new experiences and new cognitions.

3. People then try out new behaviors, both in and out of the counseling sessions.
4. With the counselor's help, they reflect on these new behaviors: What do these behaviors say about their self-definitions?
5. They then redefine their selves according to their new behaviors.
6. Feedback from others, including the counselor and other members in a group counseling setting, allows them to monitor the changes.

It may be that when this sequence occurs, therapists quite often notice that the therapy is "working." Perhaps good counseling and psychotherapy experiences can be best understood as instances of heightened self-perception.

Traditional systems of counseling and psychotherapy explain why their particular approaches work, and all offer guidelines for the process. But what is the essence of the change process? Perhaps the simplest explanation is that people change as they try new behaviors, observe these, and thus change their definitions of who they are.

Here is one example. Recovering alcoholics do not just see themselves as "alcoholics." They see themselves as "recovering alcoholics." But, recovering alcoholics who relapse no longer see themselves as recovering persons; they now see themselves as hopeless alcoholics. They have observed their own (relapsing) behavior and concluded that they will always be alcoholics. They may say something to themselves like: "I'm drinking, therefore I'm an alcoholic," and "As long as I'm an alcoholic I might as well drink and have a good time." Then they usually drink some more. Marlatt & George (1984) have called this the "abstinence violation effect." Although cognitive dissonance may also explain relapse behavior, the important point is that the cognitions, which determine self-definitions, emerge largely from observing one's own behaviors.

We are much more likely to come to know ourselves, according to self-perception theory, by observing our own behaviors in a variety of situations than by introspecting about "Who am I?" In counseling and psychotherapy, this means that the person needs to first try out new behaviors and then to reflect on what these mean to his or her self-definition.

Many traditional psychotherapy approaches are well explained from this perspective. Here are a few examples from the most prominent approaches. Rogerian counselors believe that exploring one's feelings and experience leads to greater acceptance of these and to a redefinition of the self. Rogers saw counseling as a process of "learning of self" and rejecting older ways of self-definition (Rogers & Stevens, 1967). From a self-definition and self-perception perspective: As we express our own feelings and experiences and then take a look at them, we get to know who we are.

Traditional cognitive behaviorists, such as Ellis (2000), are directive: they ask their clients to dispute and attack their own "irrational" beliefs and they routinely assign homework between sessions. From the perspective of self-perception theory, clients are taught new ways of observing their own thoughts and behaviors. Homework is crucial. It is an opportunity for clients to try out new behaviors and to consider that they are the kind of person who can do these (homework) things. Successful change occurs when clients try new thoughts and behaviors, observe them, and change their perceptions and definitions of themselves.

How we get clients to try new behaviors is a function of how creative we are and how persuasive we are. For example, Milton Erickson (Haley, 1973) would get his clients to try new behaviors in various creative ways, including having them go into a "trance." Once, when a married couple came to him for counseling, he asked them to climb a mountain outside of Phoenix. Furthermore, he insisted that their task was not completed until they had gone through a process of making attributions. He wanted them to consider why they had done this unusual assignment. What did their behavior mean? He was actively encouraging an attribution process that directly affected his clients' definitions of themselves and promoted a process of self-perception.

Psychodynamically-oriented therapy may also be most effective when it fosters self-perception. Freud (1955) believed that cure follows insight. But perhaps change is catalyzed only in those cases when the insight is based on the patient's trying out some new behavior as a result of therapy experiences.

When self-perception theory is systematically applied to the process of self-definition that occurs in therapy, several hypotheses for future empirical validation become apparent. Experiential, active approaches are more effective than passive ones. Active homework assignments with an in-session emphasis on what those homework behaviors mean to the person's definition of oneself are more effective than homework without the self definition (or no homework at all). Change should follow those insights that are based on trying new behaviors that can be self-observed rather than those based on understanding only. Counseling and psychotherapy should include a heavy dose of trying new behaviors that can be self-observed and considered in terms of self-definition.

Future research should test these hypotheses. First, the degree to which therapists recognize self-definition as a therapeutic process in their work needs to be measured. Second, clients' perceptions of key aspects of their therapy experiences need to be addressed from this perspective.

In particular, two measures beg development: a general measure of individuals' recognition of their self-perception processes and a measure of psychotherapy clients' judgments about the key change processes in psychotherapy. The degree of change and satisfaction reported by clients of therapists from different approaches can, for example, provide some answers to questions that long been waiting to be answered.

REFERENCES

Baumeister, R. F. (1982). The self. In D. T. Dilbert, S. T. Fiske, & G. Lindzay (Eds.), *The handbook of social psychology: Vol 2* (4th ed., pp. 680-740). Boston: McGraw-Hill.

Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6, pp. 1-62). New York: Academic Press.

Ellis, A. (2000). Rational emotive behavior therapy. In R. J. Corsini & D. Wedding (Eds.), *Current Psychotherapies* (6th ed., pp.168-204). Itasca, IL: F. E. Peacock Publishers.

Etchison, M. & Kleist, D.M. (2000). Review of narrative therapy: Research and utility. *Family Journal, 8* (1), 61-66.

Fazio, R. H., Effrein, E. A. & Falender, V. J. (1981). Self-perceptions following social interaction. *Journal of Personality and Social Psychology, 41*, 232-242.

Feixas, G. (1990). Personal construct theory and systemic therapies: Parallel or convergent trends? *Journal of Marital and Family Therapy, 16* (1), 1-23.

Franklin, C. & Nurius, P.S. (1996). Constructivist therapy: New directions in social work practice. *Families in Society, 77* (6), 323-327.

Freedman, J. & Combs, G. (1996). *Narrative therapy: The social construction of preferred realities.* New York: Norton.

Freud, S. (1955). *The standard edition of the complete psychological works of Sigmund Freud.* London: Hogarth Press: Vol. II Studies on hysteria.

Granvold, D.K. (1996). Constructivist psychotherapy. *Families in Society, 77*, (6), 345-357.

Haley, J. (1973). *Uncommon therapy.* New York: Norton.

Kelly, A. E. (2000). Helping construct desirable identities: A self-presentational view of psychotherapy. *Psychological Bulletin, 126* (4), 475-494.

Linden, J. A. (1994). Gratification and provision in psychoanalysis: Should we get rid of the "rule of abstinence?" *Psychoanalytic Dialogues, 4* 549-582.

Lyddon, W. J. (1990). First- and second-order change: Implications for rationalist and constructivist cognitive therapies. *Journal of Counseling and Development, 69* (2), 122-130.

Marlatt, G. A. & George, W. H. (1984). Relapse prevention: Introduction and overview of the model. *British Journal of Addiction, 79*, 261-275.

McGuire, W. J., & McGuire, C. V. (1988). Content and process in the experience of self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol.21, pp. 97-144). New York: Academic Press.

Neimeyer, R. A. (1995). Constructivist psychotherapies: Features, foundations, and future directions. In R. A. Neimeyer & M. J. Mahoney (Eds.), *Constructivism in psychotherapy* (pp. 11-38). Washington, DC: American Psychological Association.

Rogers, C. R. (1951). *Client-centered therapy*. Boston: Houghton Mifflin.

Rogers, C. R., & Stevens, B. (1967). *Person to person: The problem of being human*. New York: Simon & Schuster.

Rosenberg, M. (1986). *Conceiving the self* (2nd ed.). Malibar, FL: Krieger.

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173-220). New York: Academic Press.

Schlenker, B. R., Dlugolecki, D. W. & Doherty, K. (1994). The impact of self-presentations on self-appraisals and behavior: The power of public commitment. *Personality and Social Psychology Bulletin, 20*, 20-33.

Schwartz, R. C. (1999). Narrative therapy expands and contracts family therapy's horizons. *Journal of Marital and Family Therapy, 25* (2), 263-267.

The Role of Part-Set Cuing in the Recall of Chess Positions: Influence of Chunking in Memory.

Charles J. Huffman
James Madison University

T. Darin Matthews
The Citadel

Phillip E. Gagne'
James Madison University

A surprising finding in memory research is that cues consisting of part of the material studied can inhibit recall of the remainder of the material. However, Watkins, Schwartz, and Lane (1984) found that providing chessplayers during recall with part of a previously studied chess position did *not* inhibit recall of the remainder of the position. Chess pieces were chosen at random to be in the cued condition, however, which did not utilize how players organize information in "chunks" (patterns of related pieces). The present study used chunks to be more sensitive to how memory organization affects the occurrence of part-set cuing inhibition. Realistic game positions were presented to 12 expert and 12 novice chessplayers and recall of correct piece placement in cued and uncued conditions was measured. Consistent with previous research on chess expertise, experts recalled significantly more pieces than novices. As in the Watkins et al. study, recall was not inhibited in the cued condition. This is consistent with findings that part-set cuing is less likely to have an effect on automatic than on conscious tasks. Development of chess skill involves automatization of processes involved in recalling piece placement, so that conscious resources can be used for move selection.

Typically, memory research has shown that an event is more likely to be remembered if stimuli that were present during that event are presented again during recall. These stimuli presumably serve as contextual cues that facilitate memory, possibly by re-establishing associations made during the event. Numerous researchers (e.g., Morris, Bransford and Franks, 1977; Tulving and Pearlstone, 1966; Tulving and Thomson, 1971) have provided evidence of the facilitative effects of cues on recall. One effect, however, of cues on recall has received less attention -- specifically, the *inhibiting* influence that cues can have on recall when the cue consists of part of a previously studied set of items. This negative effect of cuing has been aptly named part-set cuing

---

Author info: Correspondence should be sent to Charles J. Huffman, Department of Psychology, MSC 7401, James Madison University, Harrisonburg, VA 22807, or email huffmacj@jmu.edu.

North American Journal of Psychology, 2001, Vol. 3, No. 3, 535-542.
© NAJP

inhibition (Basden, Basden, Church, & Beaupre, 1991; Nickerson, 1984; Peynircioglu, 1987; cf. Slamecka, 1968).

During the part-set cuing procedure, participants are provided during recall with a subset, usually chosen at random, of the items originally studied. Participants are instructed to use this subset of items as cues to help them remember the remainder of the items that were studied. When performance in the part-set cuing condition is compared to a situation in which participants are not provided with cues, recall is often found to be worse in the cued condition than in the uncued condition (Roediger, 1974; Rundus, 1973; Slamecka, 1968). This effect has been demonstrated with both related (Roediger, 1973; Watkins, 1975) and unrelated word lists (Slamecka, 1968). Furthermore, this inhibitory effect of retrieval cues on memory has also been found with nonverbal materials such as picture identification (Peynircioglu, 1987). These findings contrast sharply with studies that demonstrate the facilitative effects of retrieval cues on memory.

Given that cues can either facilitate (e.g., Tulving & Thomson, 1971) or inhibit recall (e.g., Nickerson, 1984), factors that affect the occurrence of facilitation or inhibition need to be delineated. This study examines the way in which chessplaying expertise and cuing may interact to affect recall of previously seen chess positions. The game of chess has often been useful as a laboratory task to explore how organization in memory changes as expertise develops, and how that organization affects performance on various memory tasks (see Ericsson & Smith, 1991 for a review).

In a previous study that examined the effects of part-set cuing on the recall of chess positions (Watkins, Schwartz, & Lane, 1984), chessplayers were presented with 10 board positions chosen from actual games and were asked to reconstruct the position after each presentation. For the cued condition, half of the pieces (the cue set) were already in place on the board when the participant began placing the rest of the pieces (the critical set). For the uncued condition, the board was empty of pieces when the participant began reconstruction. Correct placement of the pieces in the critical set was measured and compared between the cued and uncued conditions. Results indicated that part-set cuing neither facilitated nor inhibited recall of the remainder of the pieces. It was concluded that part-set cuing is not a sensitive test of memory organization.

In that study, however, the pieces chosen to be in the cue set were determined at random. It is possible that the way chessplayers organize information into units or "chunks" (patterns or configurations of pieces according to how those pieces interact) was not utilized, and so an important aspect of memory organization was not available to affect recall of the critical set. Research on memory conducted with chess players has consistently found that experts remember more information concerning the position of pieces on a chessboard than novices, but only when the positions resemble actual games (Chase & Simon, 1973). When pieces are placed on the board at random, experts do no better than novices. Actual game positions provide experts with highly organized information that facilitates recall, and that organization is unlikely to occur with random positions. The importance of chunking in memory for chess-specific information has been established by numerous studies (e.g., Chase & Simon, 1973; Gobet & Simon, 1996; Holding, 1992; Simon & Gilmartin, 1973). Therefore, this study was designed to be more sensitive to how part-set cuing could be affected by memory organization, by using chunks in the cue set, rather than choosing the cue set pieces at random as in Watkins et al. (1984). For example, if two pieces work together (such as a queen and bishop both aimed at the opponent's king), these pieces are likely to be organized in memory together as a chunk. In this study, these pieces both would be chosen as part of the cue set or critical set, rather than allowing one to be in the cue set and the other to be in the critical set, which could happen when random selection of pieces is used.

METHOD

Design

The design was a 2 X 2 factorial with level of expertise (experienced or novice) manipulated between-subjects and cuing (cued or uncued) manipulated within-subjects. Level of expertise was manipulated to determine if the part-set cuing procedure has an effect only on skilled, experienced players, who presumably have chess-specific information better organized than less skilled players. Watkins, et al. (1984) incorporated a second experiment with more experienced players to test this possibility. The present study included both a novice group and an experienced group in the same experiment in order to take into account all of the variance in the analyses and detect a possible interaction between the two variables of expertise and cuing. Including level of expertise in this experiment also provided a means for replicating previous studies of chess expertise and memory in which experts recalled more pieces in game positions than novices (e.g., Chase & Simon, 1973).

Participants

Twenty-four male chessplayers of various ages participated in the experiment. Participants were either university students and/or members of a community chess club. Half had experience playing in chess tournaments (the experienced group), and half had not played in tournaments (the novice group). Those participants in the novice group had some playing experience, however, and knew enough basic rules and strategies of the game to participate in this experiment. Those in the experienced group had experience playing in rated chess tournaments and, therefore, had skill ratings assigned by the United States Chess Federation (USCF). Their mean USCF rating was 1627 (see Elo, 1978 for an explanation of the USCF rating system). Although not all participants in the experienced group were experts by USCF standards (a rating of 2000 is required for expert status), two were experts and all had ratings above the national mean of 1184 (United States Chess Federation, 1999). All would be expected to regularly defeat players in the novice group. Therefore, participants in the experienced and novice groups differed in expertise as well as experience.

Materials

The same chess positions were used in this experiment that were used in Watkins, et al.'s (1984) study. The positions were from the first 10 games described in an issue of the magazine "Chess Life and Review" published by the United States Chess Federation (1975, May). These games were played by masters (USCF rating of 2200 or above). For each of these games, the first position in which 24 of the original 32 pieces remained on the board was used. The positions were therefore realistic, sensible middle game placements of pieces. For each position, instead of *randomly* assigning the 24 pieces to be in either the critical or cue set (12 in each set), as was done by Watkins et al. (1984), the 24 pieces were first divided into 12 "chunks" or pairs of pieces that cooperate in attacking the opponent's pieces or in defending their own. Typical chunk configurations consisted of king and rook in a castled position, queen and bishop both on the same diagonal aimed at the opponent's king, two rooks defending each other and controlling an open file, etc. These 12 *chunks* were then randomly placed in either the cue set or critical set (6 chunks per set). In this way the chunks were not split by allowing one piece in a chunk to be placed in the cue set and the other in the critical set, which can happen when random selection of pieces for each set is used.

Procedure

Participants were tested individually and were instructed to memorize and then reconstruct the position of the pieces for each of the 10 positions. They were told just to memorize the position of the pieces, not to generate possible moves. The 10 positions were presented in the same order to all participants. Each position was shown for 30 seconds, then removed from view and the participant immediately tried to reconstruct the position just shown using another identical chessboard and pieces. On alternate positions, the response board contained half of the pieces already in place (the cue set), and participants tried to place the remaining pieces (the critical set). This was the cued condition. For the other positions, the response board was empty and participants tried to place all of the pieces. This was the uncued condition. For half of the participants, the odd numbered positions were presented in the cued condition and the even numbered positions were presented in the uncued condition. For the other half of the participants, the converse ordering was used. In both conditions, participants were given a complete set of 32 pieces, so that they had to choose the pieces as well as place them. There was no time limit to place the pieces, but for each position all participants indicated they were finished placing them within four minutes.

RESULTS AND DISCUSSION

For each analysis the distribution normality assumption was tested using the Kolmogorov-Smirnov procedure, and in each case the results indicated that the population was normally distributed, and therefore parametric statistical tests were used. Of main interest was whether more of the critical set of 12 pieces would be recalled in the cued condition than in the uncued condition. Therefore, the dependent measure for this analysis was the number of pieces correctly recalled in the critical set: the set of pieces not already placed on the response boards in the cued condition. See Table 1 for the results. As expected, there was a main effect of expertise, $F_{(1,22)}=9.83$, $p<.01$, with the more experienced players recalling significantly more pieces correctly than the novices. There was no main effect of cuing, $F_{(1,22)}=2.02$, $p>.05$, nor an interaction between expertise and cuing, $F_{(1,22)}<1$.

The finding that the players in the experienced group recalled more pieces correctly than the players in the novice group not only substantiates the expertise manipulation, but replicates previous studies in which expert players recalled more pieces from actual game positions than novices (e.g., Chase & Simon, 1973). Although not relevant to the comparison between the cued and uncued conditions of recall of critical set pieces, the experienced group in the present study also recalled more cue set pieces in the uncued condition than the novice group, $t=3.31$, $p<.01$.

The lack of a significant difference between the cued and uncued conditions replicated Watson, et al. (1984) by finding that recall of the critical set was not facilitated, nor inhibited, by presenting the cue set during recall. This finding indicates that recall is not affected even when a method that is sensitive to how chess players organize information in memory is used. This is consistent with other studies (Basden et al., 1991; Matthews, 1995) that indicate that part-set cue inhibition is less likely to be found in highly automatized tasks than in controlled, conscious tasks.

TABLE 1 Mean Number of Critical Set Pieces per Position Placed Correctly (maximum=12) as a Function of Cuing Condition and Level of Expertise

	Cuing Condition		
Level of Expertise	Cued Mean (SD)	Uncued Mean (SD)	Total Mean
Novice group	6.47 (1.77)	6.23 (1.35)	6.35
Experienced group	8.55 (2.05)	7.98 (1.31)	8.27
Total Mean	7.51	7.11	

For example, in Experiment 2, Basden et al. (1991) found that no inhibition occurred with part-set cuing on an indirect (i.e., implicit or unconscious) test when performance between a standard condition and the indirect test condition was compared. This finding suggests that part-set cuing inhibition results from conscious influence. Additional support for this conclusion was provided by Matthews (1995), in which an attempt was made to link the inhibitory effect of part-set cues to either conscious or automatic processes. Using the process dissociation (Jacoby, 1991) and the independence remember/know procedures (Jacoby, Yonelinas, & Jennings, 1997), estimates of conscious memory were found to decrease in the presence of part-set cues, and the automatic estimates of memory were unaffected by the presence of part-set cues. Both dividing attention (Experiment 1) and delay (Experiment 2) reduced conscious but not automatic estimates of memory, thus strengthening the claim that conscious processes were being manipulated while automatic processes were not. In the final experiment, a comparison of the effects of part-set cues on explicit and implicit tests revealed the same pattern of performance as found in Experiments 1 and 2. That is, an inhibitory effect of part-set cuing was found with an explicit cued recall test (i.e., a test relying primarily on conscious memory processes) while such an inhibitory effect was absent on an implicit test of memory (i.e., a test relying primarily on automatic memory processes).

The recall of chess positions by chess players presumably involves a process that becomes highly automatized with experience and the development of expertise in order for the player to allocate controlled, effortful processing to the task of choosing the best moves to win the game. Consequently, if the recall of chess positions of expert players depends primarily on the relatively automatic retrieval of piece placement information, then the lack of an inhibitory effect of part-set cuing on recall is explainable. However, such an interpretation is complicated by the lack of an interaction between cuing and expertise. According to this explanation the novice group should have experienced inhibition in the cued condition, because the organizational structure present in the game and detected by the better players presumably would not be detected by the novices. The lack of such inhibition for the novices perhaps was due to the complexity of the game itself, however. Although not nearly as good at playing the game as the experts, the "novices" in any study involving chess, nevertheless, have to be good enough to perform rather complex tasks such as the memory task in this study. Consequently, their chess skill may have developed beyond the point of being susceptible to inhibition from the presence of the type of cue used in this study.

REFERENCES

Basden, B.H., Basden, D.R., Church, B.A., & Beaupre, P. (1991). Setting boundary conditions on the part-set cuing effect. *Bulletin of the Psychonomic Society, 29*(3), 213-216.

Chase, W.G., & Simon, H.A. (1973). The mind's eye in chess. In W.G. Chase (Ed.), *Visual information processing*. New York: Academic Press, 215-281.

Elo, A. (1978). *The rating of'chessplayers, past and present.* New York: Arco.

Ericsson, K.A., & Smith, J. (Eds.) (1991). *Studies of expertise: Prospects and limits.* Cambridge: Cambridge University Press.

Gobet, F., & Simon, H.A. (1996). Recall of random and distorted chess positions: Implications for the theory of expertise. *Memory and Cognition, 24*(4), 493-503.

Holding, D.H. (1992). Theories of chess skill. *Psychological Research, 54,* 10-16.

Jacoby, L.L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30,* 513-541.

Jacoby, L.L., Yonelinas, A.P., and Jennings, J. (1997). The relation between conscious and unconscious (automatic) influences: A declaration of independence. *Scientific approaches to consciousness.* New Jersey: Lawrence Earlbaum Associates.

Matthews, T.D. (1995). *Part set cuing impairs conscious but not automatic retrieval.* Unpublished doctoral dissertation, University of North Carolina at Greensboro.

Morris, C.D., Bransford, J.D., & Franks, J.J. (1977). Levels of processing versus test-appropriate strategies. *Journal of Verbal Learning and Verbal Behavior, 16,* 519-533.

Nickerson, R.S. (1984). Retrieval inhibition from part-set cuing: A persisting enigma in memory research. *Memory & Cognition, 12*(6), 531-552.

Peynircioglu, Z.F. (1987). On the generality of the part-set cuing effect: Evidence from nonmemory tasks. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 13*(3), 437-442.

Roediger, H.L., III (1973). Inhibition in recall from cuing with recall targets. *Journal of Verbal Learning and Verbal Behavior, 12,* 644-657.

Roediger, H.L., III (1974). Inhibiting effects of recall. *Memory & Cognition, 2*(2), 261-269.

Rundus, D. (1973). Negative effects of using list items as recall cues. *Journal of Verbal Learning and Verbal Behavior, 12,* 43-50.

Simon, H.A. & Gilmartin, K.J. (1973). A simulation of memory for chess positions. *Cognitive Psychology, 5,* 29-46.

Slamecka, N.J. (1968). An examination of trace storage in free recall. *Journal of Experimental Psychology, 76,* 504-513.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior, 5,* 381-391.

Tulving, E., & Thomson, D.M. (1971). Retrieval processes in recognition memory: Effects of associative context. *Journal of Experimental Psychology, 87,* 116-124.

United States Chess Federation (1975, May). *Chess Life and Review.* New York: Author.

United States Chess Federation (1999). *1999 Annual Rating List.* New York: Author.

Watkins, M.J. (1975). Inhibition in recall with extralist "cues." *Journal of Verbal Learning and Verbal Behavior, 14,* 294-303.

Watkins, M.J., Schwartz, D.R., & Lane, D.M. (1984). Does part-set cuing test for memory organization? Evidence from reconstructions of chess positions. *Canadian Journal of Psychology, 38(3),* 498-503.

Author Note: The authors wish to thank all of the participants in this study, including members of the Shenandoah Valley Chess Club in Harrisonburg, Virginia, as well as students at the University of North Carolina at Greensboro and at James Madison University.

How to Submit Papers to *NA*JP

The submission of good papers is very important to any psychology journal. We hope that you will support our views about what a psychology journal should be like by submitting papers now. We will try to mail a decision on your paper within 5 weeks of receipt by us. We will publish your paper in an average of 5 months of receipt of a final accepted version.

Submit 1 paper copy of each manuscript prepared according to APA guidelines. Also include a diskette with your manuscript or (preferably) e-mail the file to NAJP@mpinet.net. Include 2 large stamped envelopes so we can send your manuscript out for review. We encourage you to have your paper read by peers before submission.

If the final paper is accepted, please send a check or purchase order for the amount specified with your final draft (billed at $25 dollars per typeset page in units of two pages).

Submission of a manuscript to the Journal is a contract to publish with the Journal if the paper is accepted. Because of our rapid editorial feedback, simultaneous submission to other journals is not allowed. We will offer brief comments on the suitability of a topic for publication in *NA*JP before a submission.

Send all correspondence to: *North American* Journal of Psychology, 240 Harbor Dr., Winter Garden, FL 34787, phone 407-877-8364.

SUBSCRIPTIONS

All subscriptions are on a calendar year basis only. If you order late in the year, the earlier issues will be sent to you. Institutional rate, $65. Individual rates, paid by check, $30. Add $7 for Canadian subscriptions. Add $8 for subscriptions outside North America. Phone orders accepted with Visa, MasterCard or American Express. Call 407-877-8364. The *North American* Journal of Psychology is now published semi-annually in June & December.

Visa ☐ AMEX ☐ MC ☐

Card # _______________ Exp. Date _______

Name _______________________________

Address _____________________________

City _______________ State _______ Zip _______

A Summary of Our Philosophy

The ***North American*** **Journal of Psychology** publishes scientific papers that are of general interest to psychologists and other social scientists. Papers may encompass a wide variety of topics, methods, and strategies. It is our belief that many good scientific papers are rejected for invalid reasons. Some journals reject up to 90% of their submissions because they don't have space. Furthermore, studies of the review process strongly suggest that not all of the lucky 10% are better papers than those that get rejected. In our view many of those rejected papers are meritorious with need for only minor revisions. Perhaps our philosophy can be summed up this way: we are looking for ways to strengthen good papers rather than searching for weak excuses to reject them.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

☒ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").